AD-A119 076 DECISIONS AND DESIGNS INC MCLEAN VA F/G 9/5
APPROXIMATE REASONING METHODS FOR DECISION AIDS.(U)
MAY 82 P J STICHA, J F PATTERSON, J J WEISS F30602-80-C-0156
UNCLASSIFIED TR-82-6-157 RADC-TR-82-127 NL

1 OF 2
AD A
119076

[illegible]
Final Technical Report
May 1982

APPROXIMATE REASONING METHODS FOR DECISION AIDS

Decisions and Designs, Inc.

Paul J. Sticha, John F. Patterson and Jonathan J. [illegible]

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, NY [illegible]

[illegible]

[illegible]

APPROVED: John J. Maier

JOHN J. MAIER
Project Engineer

APPROVED:

A. J. DRISCOLL, Colonel, USAF
Chief, Intelligence & Reconnaissance Division

FOR THE COMMANDER:

JOHN F. [illegible]
Acting Chief, [illegible]

[illegible]

[illegible]

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER RADC-TR-82-127	2. GOVT ACCESSION NO. AD A119 076	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) APPROXIMATE REASONING METHODS FOR DECISION AIDS		5. TYPE OF REPORT & PERIOD COVERED Final Technical Report Apr 80- Mar 82
		6. PERFORMING ORG. REPORT NUMBER TR 82-6-157
7. AUTHOR(s) Paul J. Sticha John F. Patterson Jonathan J. Weiss		8. CONTRACT OR GRANT NUMBER(s) F30602-80-C-0156
9. PERFORMING ORGANIZATION NAME AND ADDRESS Decisions and Designs, Inc. Suite 600, 8400 Westpark Drive, P.O. Box 907 McLean VA 22101		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62702F 45941651
11. CONTROLLING OFFICE NAME AND ADDRESS Rome Air Development Center (IRDT) Griffiss AFB NY 13441		12. REPORT DATE May 1982
		13. NUMBER OF PAGES 136
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Same		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Same

18. SUPPLEMENTARY NOTES

RADC Project Engineer: John J. Maier (IRDT)

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Decision Theory
Decision Analysis
Decision Aids
Computer-Based Modelling
Fuzzy Set Theory
Bayes Theorem
Target Intelligence
Target Nomination
Computer Mapping
Videodiscs

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report describes a project of analytic research to develop a decision aid using methods of approximate reasoning. Approximate reasoning, in its most general sense, refers to a set of techniques that attempt to reflect the inherent imprecision in human language and thought. In a stricter sense, it refers to a specific collection of methods subsumed under the theory of fuzzy subsets.

cont next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Cont'

In the course of this research effort, a prototype decision aid was developed for Air Force target nomination. Target nomination refers to the process by which Air Force intelligence officers develop an ordered list of targets as candidates for attack. The computer-based decision aid developed combines methods from decision analysis, fuzzy set theory, and computer graphics in order to achieve a system that could facilitate an intelligence analyst's efforts.

The basic analytic approach involves two decision-analytic models, an expected value model, and a resource allocation model. Methods derived from the theory of fuzzy subsets were used to adjust model parameters in the light of user feedback, and these methods were then compared to Bayesian updating methods. The Bayesian methods were judged superior in the current model; however, the fuzzy-subsets approach offers considerable promise in implementations with sufficient computer resources.

The decision aid interacts with the user through a geographical display system. This system provides a map background onto which information about targets may be overlayed. The geographical display system uses videodisc storage and image processing techniques to provide rapid access to a wide range of map coverage.

A

SUMMARY

The research described in this report seeks to develop a decision aid that uses methods of approximate reasoning. In its most general sense, approximate reasoning refers to an attempt to reflect the inherent imprecision in human language and thought. In a stricter sense, it refers to a specific collection of methods subsumed under the theory of fuzzy subsets. A specific goal of this research is to compare a decision aid based on the theory of fuzzy subsets to one based on Bayesian methods.

In the course of this research effort, a prototype decision aid was developed for Air Force target nomination. Target nomination refers to the process by which Air Force intelligence officers develop an ordered list of targets as candidates for attack. The process is performed by a target officer in the Targets Intelligence branch of the Combat Operations Intelligence Division within the Tactical Air Control System. During wartime, target nomination is performed on a regular basis on a cycle corresponding to that of the Air Tasking Order.

The decision aid combines methods from decision analysis, fuzzy-set theory, and computer graphics in order to achieve a system that could facilitate an intelligence analyst's efforts. The basic analytic approach involves two decision-analytic models, an expected value model, and a resource allocation model. The expected value model integrates the importance of target with variables affecting the likelihood that the target will be destroyed to obtain an overall measure of value. The

resource allocation model assigns sorties to targets in order to obtain the greatest expected value for the number of sorties allocated. The decision-analytic model is organized around the decision prototype concept. This concept is an attempt to implement the more general concept of an advisory decision aid.

Methods derived from the theory of fuzzy subsets are used to adjust model parameters in the light of user feedback. These methods were compared to Bayesian updating methods. Although both methods were fundamentally successful, each encountered some problems: the Bayesian method was fast and convenient, but too simple to capture complex or subtle features of the problem. The fuzzy sets method, on the other hand, presented some implementation problems, largely due to its higher complexity. Given a system with sufficient computational abilities, the fuzzy-sets approach seems to offer greater promise.

The decision aid interacts with the user through a geographical display system. This system provides a map background onto which information about targets may be overlayed. The geographical display system uses videodisc storage and image processing techniques to provide rapid access to a wide range of map coverage. The user is provided with the abilities of scrolling or panning, zooming or scale change, controlling feature separates, and accessing photography. In addition, target data may be overlayed on the map background and colored to indicate the values of data variables. A touchscreen provides an additional method of natural communication between the user and the decision aid.

ACKNOWLEDGMENTS

The authors wish to thank a variety of individuals who made significant contributions to the progress of this project. Gen. Richard R. Stewart provided much guidance on the direction of this project in addition to the substantive expertise on which the model was based. The decision aid took shape through discussions with and the efforts of many decision analysts, including Dennis M. Buede, Michael F. O'Connor, and Cameron R. Peterson. Excellent programming efforts were provided by Audrey W. Lipps and Kenneth J. Frier under the direction of L. Scott Randall.

The development of a videodisc involved many steps and was aided by several organizations and individuals. Thanks are due to the Defense Mapping Agency, the Defense Intelligence Agency, and the Defense Advanced Research Projects Agency. Ray J. Kerns of Computer Camera Service was astute in his understanding of the filming requirements for this project, and Jan Smoot of Perceptronics, Inc. coordinated the production of the videodisc.

TABLE OF CONTENTS

TABLE OF CONTENTS (Con't.)

TABLE OF CONTENTS (Con't.)

TABLE OF CONTENTS (Con't.)

FIGURES

TABLES

APPROXIMATE REASONING METHODS
FOR DECISION AIDS

1.0 INTRODUCTION

This report describes a project of analytic research to develop a decision aid using methods of approximate reasoning. Approximate reasoning, in its most general sense, refers to a set of techniques that attempt to reflect the inherent imprecision in human language and thought. In a stricter sense, it refers to a specific collection of methods subsumed under the theory of fuzzy subsets (Zadeh, 1965). In either case, the goal is to aid decision making by constructing an aid that captures the gist of a problem, while remaining robust under the numerous sources of indeterminancy inherent within that problem.

Air Force target nomination was selected as the problem domain for this research effort. It refers to the process by which Air Force intelligence officers develop an ordered list of targets as candidates for attack. During a war, target nomination can take place daily in a complex and dynamic environment. Therefore, it is an ideal choice as a problem domain. The complexity, uncertainty, and urgency of the problem favor approximate reasoning over a more time-consuming, detailed approach. Also, the repetitiveness and importance of target nomination offer justification for the development of a computer aid whose initial expense can only be balanced by repeated applications.

1.1 The Problem Domain

Target nomination is a process by which potential targets are organized according to their value to and consistency with the commander's concept of operation. The target nomination process is principally relevant to interdiction and offensive counter air missions. The process occurs daily and forms a part of the development of the Air Tasking Order (ATO).

Target nomination is performed by a target officer (target nomination) in the Targets Intelligence branch of the Combat Operations Intelligence Division within the Tactical Air Control System. The Target Nomination Officer coordinates his activities with the Combat Plans Division and submits a nomination list to the Director of Combat Operations for inclusion in the daily ATO.

The target nomination list is an ordered list of potential targets which maximize the benefit, given the total number of sorties available to be assigned to targets, and the guidance from the commander. The nomination list also includes alternate targets which offer high benefit, should some of the nominated targets become infeasible. Development of the target nomination list is not concerned with details about assignment of specific aircraft to specific missions, nor is it concerned with other aspects of mission planning.

Target nomination occurs in a complex, dynamic, and uncertain environment, with great constraints on the time allowed for the process. In addition, the process is repetitive, occurring on a schedule corresponding to that of the ATO. Because of the complexity and the importance of target nomination, this task is an excellent candidate for building a decision aid using approximate reasoning methods.

1.1 The Problem Domain

Target nomination is a process by which potential targets are organized according to their value to and consistency with the commander's concept of operation. The target nomination process is principally relevant to interdiction and offensive counter air missions. The process occurs daily and forms a part of the development of the Air Tasking Order (ATO).

Target nomination is performed by a target officer (target nomination) in the Targets Intelligence branch of the Combat Operations Intelligence Division within the Tactical Air Control System. The Target Nomination Officer coordinates his activities with the Combat Plans Division and submits a nomination list to the Director of Combat Operations for inclusion in the daily ATO.

The target nomination list is an ordered list of potential targets which maximize the benefit, given the total number of sorties available to be assigned to targets, and the guidance from the commander. The nomination list also includes alternate targets which offer high benefit, should some of the nominated targets become infeasible. Development of the target nomination list is not concerned with details about assignment of specific aircraft to specific missions, nor is it concerned with other aspects of mission planning.

Target nomination occurs in a complex, dynamic, and uncertain environment, with great constraints on the time allowed for the process. In addition, the process is repetitive, occurring on a schedule corresponding to that of the ATO. Because of the complexity and the importance of target nomination, this task is an excellent candidate for building a decision aid using approximate reasoning methods.

1.2 Analysis Concepts

As decision aids and decision support systems are developed and applied to a greater number of problems, it is likely that the concepts underlying their development will become more poorly defined. In particular, because of the variety of methods in which the systems are applied, it be increasingly difficult to arrive at a single definition of a decision support system. In this circumstance, it will be helpful to be able to characterize the differences among decision support systems in order to organize the variety of systems that are available. In particular, the following two issues represent the differences among many current decision aids.

1. The extent to which the aid's recommendation is based on knowledge supplied by the user rather than by the model.

2. The extent to which the aid's recommendation is viewed as a final answer to the problem rather than as an organizing framework to help the user determine the final answer.

The source of the information on which the aid makes a recommendation has a great effect on the way the aid will interact with the user. In an aid that obtains information largely from the user, the bulk of the interaction between the user and the aid will consist of procedures in which the aid interrogates the user in order to assess model parameters. Effort must be made to make the assessment procedure as efficient as possible. Then, if assessments are made accurately and the aid contains a reasonable model, it would be expected that the results would be close to optimal, although high precision would not be possible due to the inherent inaccuracy of human judgments.

An aid that has considerable knowledge of its own, and that makes a recommendation with little user input, can contain a detailed model of the problem that leads to a precise answer. In situations in which the assumptions of the model are correct, this answer will be superior to that of an aid based on the judgments of the user. On the other hand, if the assumptions are incorrect, there is a chance of extreme error in the model results.

The above discussion suggests that some user inputs are necessary in many decision aids, especially for those in which the situation is not well-defined. In addition, the user should be given the opportunity to check the basic assumptions of the aid to determine if they are met by the problem at hand. The aid for targeting assesses key assumptions about mission objective and user preferences from the user, while it calculates the expected damage primarily from its own knowledge of the situation.

There are several ways in which a decision support system may be said to aid a decision. At one extreme, the system supplies information or an organizing framework to aid the user in making his own judgment. At the other extreme, the system itself makes the final judgment regarding the problem. The approach taken in this project is between these two extremes. It is based on the concept of the Advisory Decision Aid (Patterson, Randall, and Stewart, 1981). In this approach, the decision aid is viewed as an advisor that may suggest solutions to problems, and may evaluate solutions generated by the user. The decision aid produces a "straw man" solution to the problem, but the aid is responsive to user feedback, and will adjust the initial solution to reflect user criticisms.

Thus, the current aid combines information from the user with its own knowledge to produce a recommended target nomination list. The information from the user is chiefly concerned with objectives and preferences; the information from the aid is chiefly concerned with target size, expected damage, and general knowledge about targeting. Because the model operates under the general guidance of the user, it is expected that it will produce a solution that is generally correct. However, the model will undoubtedly be incorrect in several of its details. Consequently, the decision aid includes procedures that allow it to respond to user criticism in order to produce a final solution. Details of the modeling approach are presented in Section 3.0. The mechanism for adjustment in response to user feedback is presented in Section 4.0.

1.3 Approximate Reasoning

The terms "approximate reasoning" and "fuzzy sets" have become ambiguous due to their use in a wide variety of contexts, often with different meanings. This paper specifies two definitions of approximate reasoning, a weak definition and a strong definition. The weak definition represents some intuitions about the relationship between language, concepts, and real numbers that are relatively uncontroversial and widely accepted. The strong definition represents a precisely-stated theory of approximate reasoning which may be put to a logical or empirical test.

The weak definition of approximate reasoning reflects the feeling that category boundaries are not precise. Thus, it is not possible to find the exact age at which a young person becomes old, the exact height at which a short person becomes tall, the exact weight at which a thin person becomes fat, and so forth. The imprecision of category boundaries comes from the fact that words, and the concepts they represent, do not have precise numerical representation. Thus, although the ages of individuals varies more or less continuously, there are only a limited number of language concepts with which to communicate and understand the age of an individual. Consequently, each concept must stand for a range of ages. Moreover, the concepts have imprecise meanings with respect to the underlying continuous variables of interest.

A second intuition on which the weak definition of approximate reasoning is based is that human judgments are not precise. There is always some margin of error in any assessment process. Consequently, an analytical method which requires or assumes precise inputs will always give results that are in error.

An approximate reasoning method under this weak definition is one which recognizes sources of imprecision in concept definitions and parameter assessments. Under this definition, many analytical methods may be classified as using approximate reasoning. Examples include probability theory (because it recognizes imprecision in the prediction of future events) and utility theory (because it recognizes the imprecise boundary between preferred and not-preferred outcomes), as well as the theory of fuzzy sets (in the strong sense). In fact, almost any method that uses continuous variables may be thought of as an approximate reasoning method in some sense of the term.

It is easy to see that the weak definition of approximate reasoning offers little guidance on the development of specific analytical methods. It does, however, capture the spirit of approximate reasoning. Consequently, the general operation of the aid was designed to reflect a sensitivity to imprecision wherever possible. Specifically, imprecise inferences are made from inputs of user preferences and feedback. In addition, analysis results are displayed using color scales on a map background display.

The strong definition of approximate reasoning is captured in the formal methods of the theory of fuzzy subsets (Zadeh, 1965). If X is a set, then a fuzzy subset, S, of X is defined by a function, μ_S, that assigns a number in the closed interval from 0 to 1 to each element of X. *The function*, μ_S, is termed the membership function of S and describes the extent to which each element of X is a member of S.

Set theoretic concepts of complement, union, and intersection have been generalized to apply to fuzzy subsets of a set. Specific examples follow:

- o The membership function of the complement of a fuzzy subset, S, has a value equal to 1 minus the membership function of S. For example, if T is the complement of S and $\mu_S(x) = .7$, then $\mu_T(x) = .3$.

- o The membership function of the union of fuzzy subsets, S and T, has a value equal to the larger of the two component membership functions. For example, if $\mu_S(x) = .7$ and $\mu_T(x) = .4$, then $\mu_{S \cup T}(x) = .7$.

- o The membership function of the intersection of fuzzy subsets, S and T, has a value equal to the smaller of the two component truth values. For example, if $\mu_S(x) = .7$ and $\mu_T(x) = .4$, then $\mu_{S \cap T}(x) = .4$.

Several alternate definitions for these set-theoretic functions have been proposed, and there is currently some controversy over what the best form of these functions should be. The current project does not address this controversy, and uses the above rules in its formal definition of approximate reasoning.

The strong definition offers specific guidance regarding the development of analytical methods. The formal methods prescribed by the theory of fuzzy subsets were used to develop a method by which the decision aid would respond to user feedback. This method was compared to a probabilistic method based on Bayesian adjustment of probabilities. The adjustment methods are described in Section 4.0.

1.4 Overview of the Target Nomination Aid

As a result of this project, Decisions and Designs, Inc. (DDI) has developed a prototype target nomination aid. The aid combines methods from decision analysis, fuzzy set theory, and computer graphics in order to achieve a system that could facilitate an intelligence analyst's efforts.

Figure 1-1 depicts the general design of the aid. It consists of two primary subsystems: a geographical subsystem and an analysis subsystem. The analysis subsystem consists of three parts: value analysis, sortie allocation, and user feedback adjustments. Each of these four main components of the aid embodies a different principle for constructing user-friendly decision aids.

1. Information must be presented in context and in the manner in which a user is accustomed to understanding it (geographical display system).

2. Information related to the decision should be transformed into a representation that is both useful to the computer and meaningful to the decision maker (value analysis).

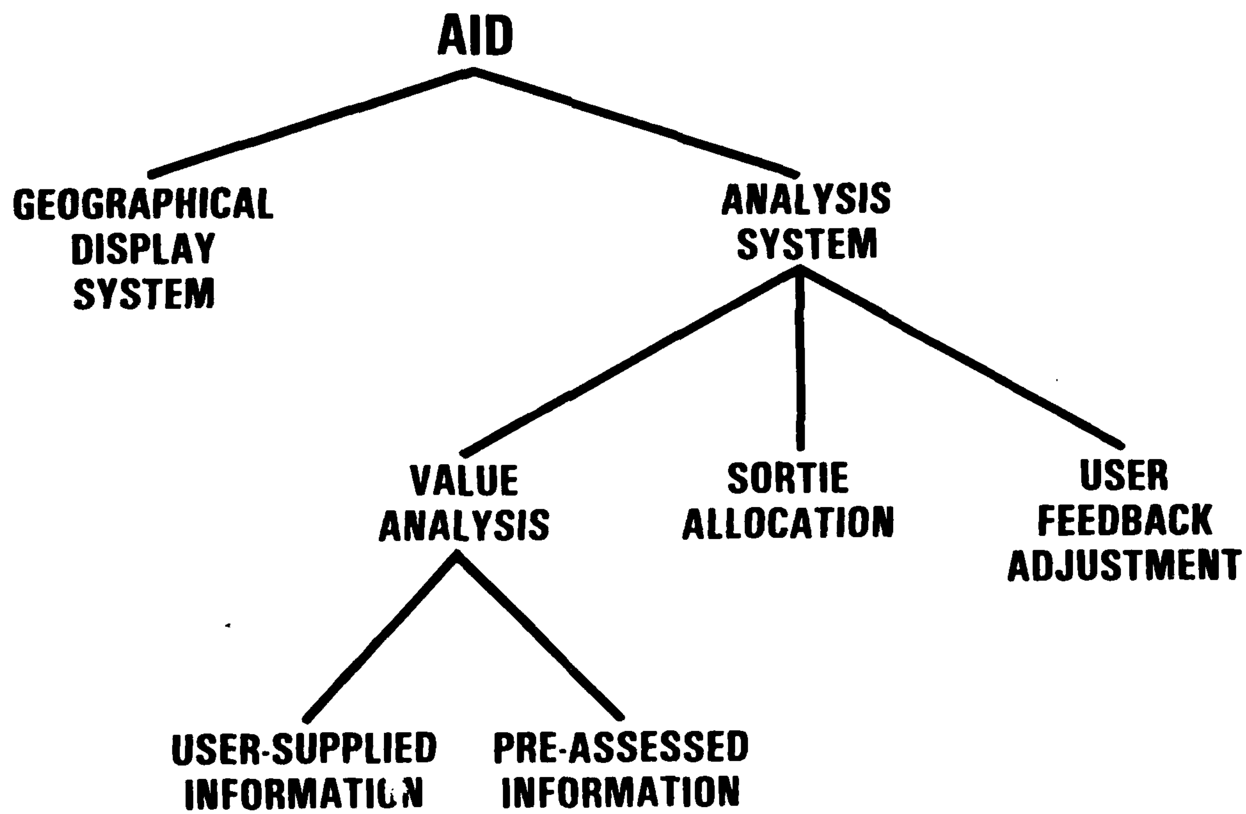

Figure 1-1

COMPONENTS OF TARGETING AID

3. An analyst's time can be better spent modifying a candidate solution generated by the computer than by simply reviewing the raw data (sortie allocation).

4. A computer-generated solution must be capable of user modification so that it can capture information and knowledge not present in either its data base or program (user feedback adjustment).

The sections that follow describe how the prototype target nomination aid has sought to achieve these four goals. Specifically, Section 2.0 describes the geographical display subsystem; Section 3.0 describes the value analysis and sortie allocation subsystems; and Section 4.0 describes the user feedback adjustment subsystem. Finally, Section 5.0 draws conclusions and gives recommendations on the future use of these methods for decision aids.

2.0 THE GEOGRAPHICAL DISPLAY SYSTEM

One of the most critical aspects of any computer aid is its interface with the user. This interface must be fast, easy to use, and meaningful. If cumbersome or difficult to learn, the aid can become useless or even detrimental to the organization it is intended to help. All too frequently, the result is an aid that is computationally correct, but ignored, because the aid's outputs cannot justify the effort required to obtain them.

For military problems, in general, and for Air Force target nomination, in particular, the best means for presenting computer information is a map. Not only is the user quite familiar with this mode of presentation, but his problems are often inherently spatial. For example, target nomination requires the user to understand the capabilities of many interacting enemy facilities and combat units. Since the character of these interactions frequently depends on the physical distance between the units, plotting targets by location is essential. Also, it is necessary to plot the targets in relation to the context within which they operate, e.g., the roads, railroads, rivers, terrain features, and other geographic features.

Of course, the importance of maps is assumed by military planners, but not necessarily by computer aid designers. Consider a target data base with geographical information represented in terms of latitude and longitude, but lacking the capacity to plot information upon a map. In all likelihood, the analyst who builds the data base extracts his information from maps and aerial photography, while the planner who uses it plots the information back onto a map in order to see the target relationships. Forgetting the value of the data base as a storage and distribution medium, it seems more a nuisance than an aid. It requires pictorial information to be translated by the user into an alphanumeric code, which is translated by another user back into a picture.

The alternative to this situation is to attach a mapping capability to the data base. This permits the analyst to enter information by pointing to locations on the map display. Also, it permits the planner to examine the information by direct display upon a map. The translations from picture to alphanumeric code, and back to picture no longer encumber the user, but are accomplished, instead, by the computer.

In all fairness to computer aid designers, the absence of mapping capability has not been due to a failure to appreciate its value. Indeed, mapping systems have existed for many years. They are, however, austere. They lack detail and shading; often they offer only small areas of coverage; and they are slow. Until recently, the technology has not been available to construct a rich computer mapping capability.

The primary technological changes that permit improved computer mapping are the development of optical videodiscs and frame buffers. Maps require a large storage capacity, especially if they are to resemble paper maps. Optical disks, whether videodiscs or digital disks (see Goldstein, 1982 for a discussion of the distinction), provide the required low cost, high density storage medium. Map usage also requires the means to overlay computer accessed information upon the map display; this is facilitated by the use of frame buffers. Together, these technological advances render a geographical display system increasingly feasible.

The present section discusses the videodisc mapping capability incorporated as part of the target nomination aid. Building upon the work at the Machine Architecture Group at M.I.T. (Negroponte, 1979), DDI originally developed its mapping capability on behalf of the Defense Advanced Research Projects Agency (see Patterson, et al., 1981). The results of this effort were then modified and enhanced for inclusion as part of the target nomination aid. The three sections that follow present the system's capabilities for both mapping and map overlays, as well as discuss the procedure for acquiring maps. The hardware configuration used by DDI is discussed in Appendix A.

2.1 The Computer Atlas

At the heart of the geographic display system is a computer atlas. In computer jargon one might refer to it as a capability for "soft copy maps." In either case, the idea is to provide rapid access to a large collection of maps.

If one were to literally mimic a paper atlas, then a computer atlas would be little more than a filing system with an index page. This is, in fact, one approach to the storage of maps on a videodisc. The videodisc can store 54,000 video images or frames on one side of a disc. Each of these is independently accessible by simply typing the desired frame number into the videodisc controller. Then, by using the first few frames as an index and equating map regions with frame numbers, a videodisc player becomes an atlas.

Fortunately, the optical videodisc provides the capability to construct a much better atlas. Considerable improvement is provided by linking the videodisc to a computer and storing the index in a computer accessable form. Given user input concerning the desired geographical region the computer can now determine the frame number based on the index, and issue the necessary command to the videodisc. The frame numbers, although critical to the operation of the system, are now invisible to the user.

This is, in essence, the primary value of a computer atlas, i.e., it hides the page numbering. Because of this, the computer atlas is much easier to use than a paper atlas and provides capabilities not available with paper. DDI has investigated four such capabilities:

1. Scrolling or panning

2. Zooming or scale change

3. Control of feature separates

4. Access to photography

2.1.1 Scrolling or panning - One of the more annoying aspects of a paper atlas is the fact that not all maps of adjacent regions can be adjacent within the atlas. The map of Colorado can be adjacent to Kansas and Nebraska, but then it is likely to be several pages away from Utah and New Mexico. No matter how the pages of the atlas are ordered, some adjacent maps will be difficult to find without skipping several pages.

This problem is a result of storing information that is organized in terms of two dimensions within a one dimensional filing system. In a book, the only way to overcome its one dimensional nature is to record page numbers at the edge of each map, thereby indicating the location of any adjacent map. With a videodisc system a similar technique can be used but without requiring the user to actually access the adjacent map. Instead, he simply indicates his desire to see the map to the north or the map to the south and the computer identifies which map is appropriate. Since this happens rapidly, from the user's point of view it is as if he is looking at a window onto a large "virtual" map and the window shifts north, south, or in any direction that he indicates. This capability is known as "scrolling" among computer scientists and "panning" among photographers. Here, it will be referred to as "scrolling."

DDI's geographical display system actually contains two approaches to scrolling: discrete and continuous. For discrete scrolling, the maps are stored as in Figure 2-1, with some overlap. Then, whenever the user requests an adjacent map by deflecting a joystick, the videodisc searches for the requested map. When this map is found, it suddenly replaces the old display, thereby jumping the user's window in the requested direction.

Despite the appearance of Figure 2-1, the actual implementation used an overlap of 67 percent for adjacent frames. Although studies by Moses and Maisano (1979) indicate that 25 percent overlap is sufficient for discrete scrolling, the optimal amount of overlap is actually an open question, since it is likely to vary as a function of access time, map complexity, and task. DDI's choice of 67 percent was dictated more by the requirements of continuous scrolling than by any empirical evidence.

Continuous scrolling is problematic with a videodisc. The difficulty is due to the slow random access of frames (1-3 seconds) and the large amount of overlap required to lend the appearance of continuity. Consider, for example, a system using the videodisc's play mode to simulate continuous scrolling. The play mode accesses sequential frames on the videodisc every 1/30th of a second. Thus, a sequence of thirty frames overlapped by 97 percent would be required to scroll across one frame in one second.

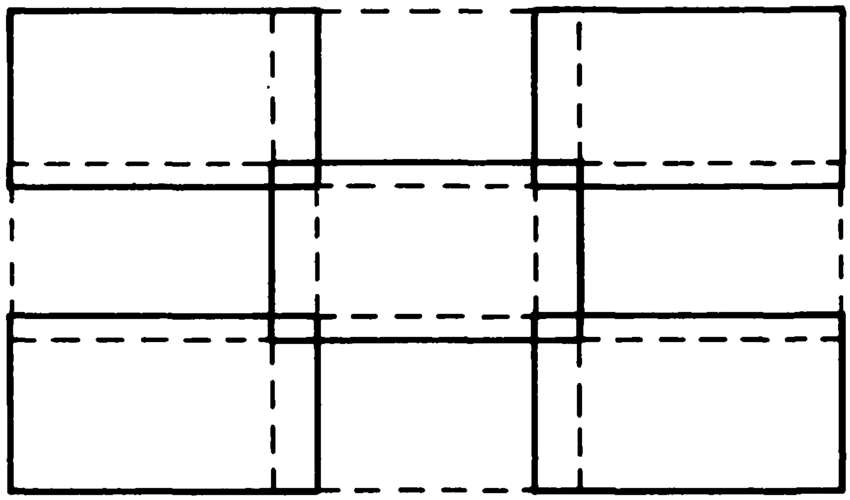

Figure 2-1
ADJACENT MAP FRAMES WITH OVERLAP

The play mode system encounters another difficulty, since it relies on the storage sequence to achieve rapid access. If, for example, map segments are stored in a sequence from east to west, what will happen when the user wants to scroll to the north or south? If, on the other hand, the map segments are stored in a north-south sequence, what will happen when the user requests east-west scrolling? Any solution to continuous scrolling which uses play mode must be prepared to duplicate storage for each direction of scrolling.

An alternative approach to continuous scrolling, and the one implemented by DDI, takes advantage of the capabilities of a frame buffer. A frame buffer is a rather large array of computer memory that can store one video frame. By using an A/D converter known as a frame grabber, the output of a videodisc player can be digitized and placed within the frame buffer. Typically, the frame buffer will contain a 512 x 512 array of memory elements, each of which can represent up to 256 intensity values. The elements of the array, known as pixels (for picture elements), are read in a scanning fashion every 1/30th of a second in order to produce a video image.

The feature of the frame buffer that facilitates continuous scrolling is the ability to alter the scanning sequence for production of the video output. One way to alter the scanning sequence will be called hardware zooming. This technique allows the frame buffer to be underscanned by a specified amount. In the case of 2:1 hardware zooming, only one quarter of the frame buffer is displayed. Figure 2-2 depicts a 2:1 hardware zoom. For the quarter frame region that is displayed, each pixel translates into a four pixel region on the display. In essence, the intensity value of each pixel has been replicated to create a block of four equal intensity pixels.

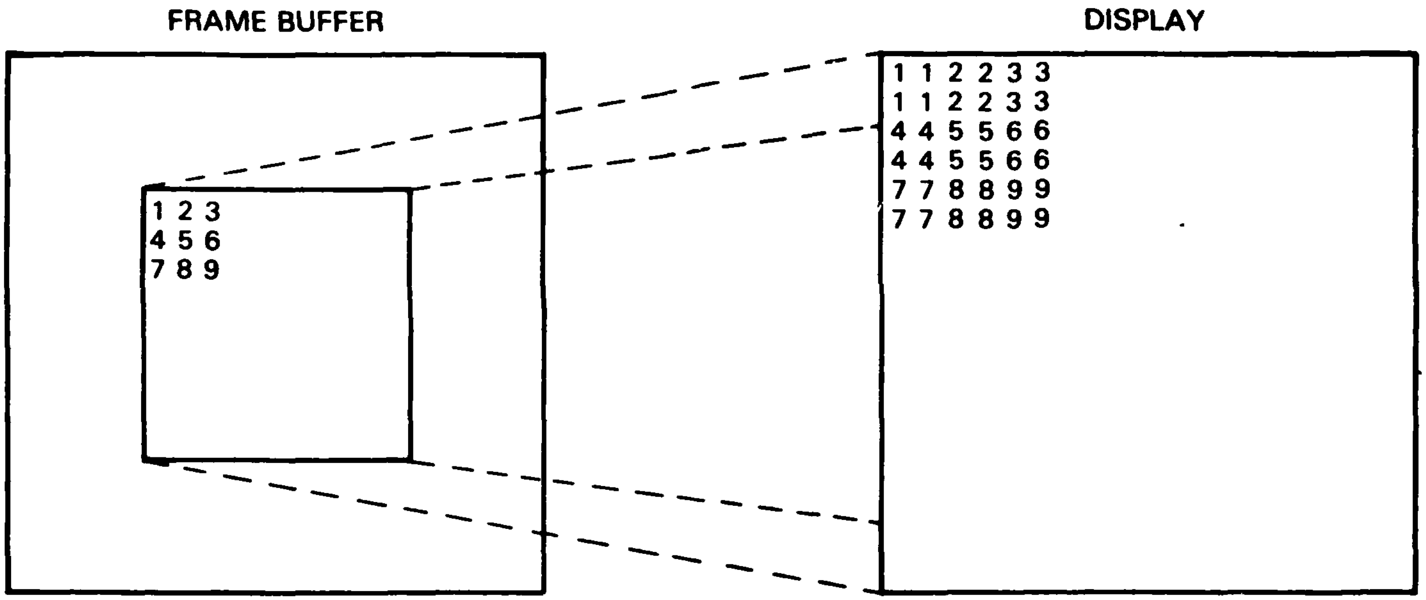

Figure 2-2

EXAMPLE OF 2:1 HARDWARE ZOOMING

The other way to alter a frame buffer's scanning sequence is to change the pixel at which the scan starts. This will be called hardware scrolling. Acting in combination with 2:1 hardware zooming, hardware scrolling implies that the quarter frame being displayed can be any contiguous quarter frame. Since the position of the quarter frame can be arbitrarily changed within 1/30th of a second, it is possible to smoothly scroll the quarter frame window across the whole frame at any desired rate.

Hardware zooming and scrolling can provide continuous scrolling within a frame, but are constrained to the frame. To achieve continuous scrolling from frame-to-frame, the videodisc player must begin searching for the adjacent frame in anticipation of its need (Figure 2-3). Then, when the hardware scrolling reaches its limit, i.e., the edge of the frame, the videodisc is ready with the adjacent frame so that a switch can occur. Since the new frame and the old frame overlap by at least 50 percent (and in this case 67 percent), one is assured that the new frame contains a quarter-frame portion that corresponds exactly with the quarter frame being abandoned on the old frame. Finally, the rate of scrolling within a frame must be limited to a rate that permits the videodisc search to be completed before the frame boundary is encountered.

This approach to continuous scrolling has been quite effective, but warrants a number of informal evaluative comments. First, it is unclear that a 67 percent overlap is needed to support this approach to scrolling. This large overlap was used to provide extra insurance against sudden alterations in direction of scroll. It helps to guarantee that upon making a switch, the new quarter frame window is relatively close to the middle of the frame and therefore has room for movement in all directions. Informal observation suggests that such rapid alterations are

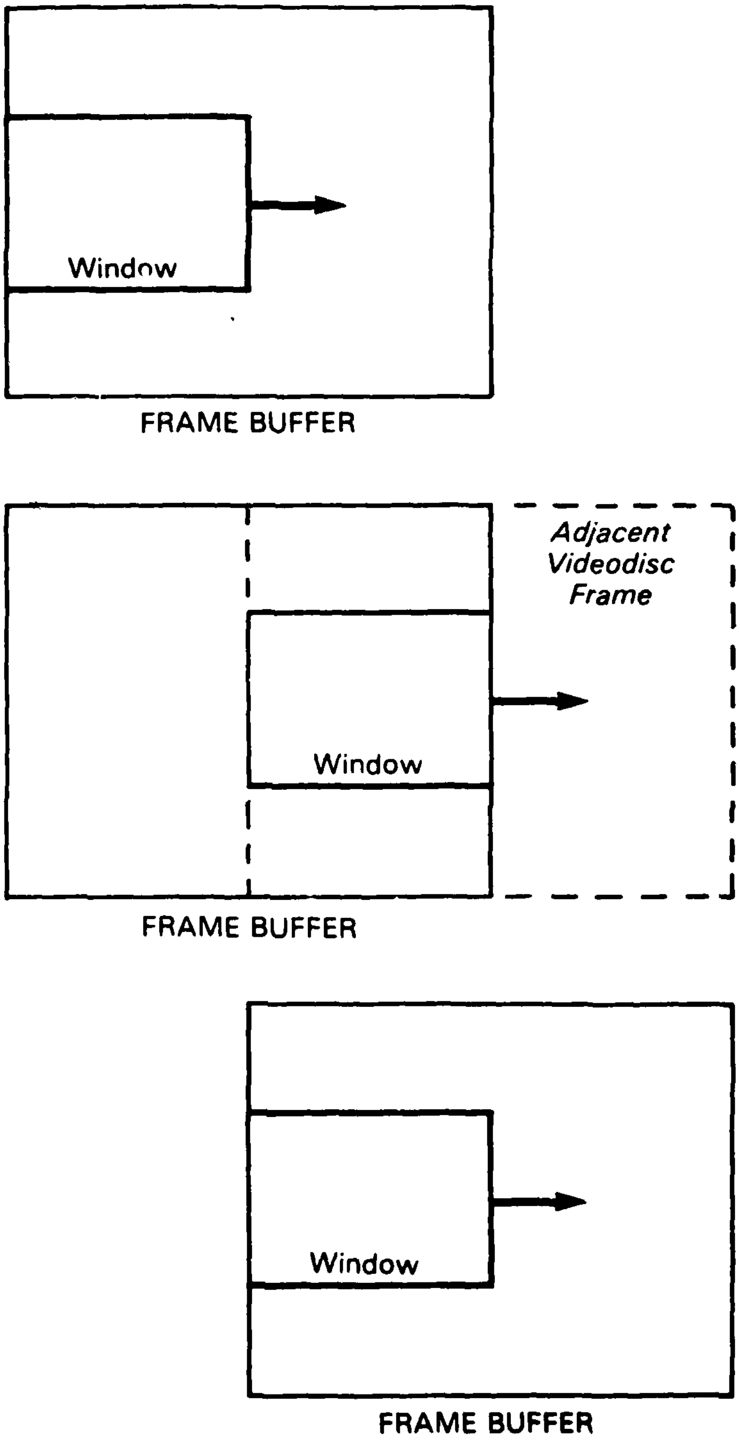

Figure 2-3

CONTINUOUS SCROLLING TECHNIQUES

rare and that when they occur the user is willing to tolerate a disruption of the smooth scrolling. Therefore, overlap much closer to 50 percent is likely to be sufficient.

Second, it is not clear that continuous scrolling is always preferable to discrete scrolling. Although it does facilitate the user's sense of continuity, it is not any faster than discrete scrolling. Also, it must sacrifice resolution and therefore detail.

Finally, continuous scrolling is clearly more expensive, both in terms of hardware and software, than discrete scrolling. Perhaps the primary situation which argues in favor of continuous scrolling is a task that requires one to track a linear geographic feature. In the absence of such a task, continuous scrolling may be unable to justify its expense and resolution loss.

2.1.2 Zooming or scale change - Zooming or scale change refers to the ability to obtain maps at progressively larger or smaller scales, but centered at the same location. In a paper atlas, this capability is usually provided by a map of the world at the front of the atlas, larger scale regional maps at the beginning of chapters, and even larger scale maps within the chapters. The computer atlas provides the same set of maps, but allows the user to quickly move from one scale to another.

DDI's actual implementation of scale changes involves two techniques. The first uses hardware zooming as described in Section 2.1.1. This can occur very rapidly and simply involves changing the display so that it shows either a quarter frame or a full frame. The second technique uses additional maps on the videodisc. (Maps at 1:1,000,000; 1:250,000; 1:50,000; and 1:12,500 were stored.) This scale change occurs slowly (1-3 sec.), since it requires a videodisc search. It does, however, display a new map with new detail that is appropriate to the changed scale.

Informal observation suggests that the present approach to scale changes could bear improvement. First, hardware zooms are unnecessary unless continuous scrolling is used. Second, discrete zooming is disorienting especially because the maps change both in detail and scale. Finally, in contrast to the present implementation, the display system should permit the user to skip intermediate scales rather than being compelled to climb progressively through the scales.

2.1.3 Control of features separates - One aspect of paper maps that becomes apparent when contrasted with computer maps is their overabundance of information. For any one application, a paper map will contain much more information than is necessary. Indeed, a paper map must contain the union of the information required by each of its users. Therefore, each user will consider the map cluttered.

Anderson and Shapiro (1979) point out the capacity of computer maps to provide clutter control; they provided their users the means to control the features that were combined to form the map. Their experience suggested that users could be relied upon to control clutter in a responsible fashion.

The geographical display system implemented by DDI provides the means to control clutter. Figure 2-4 demonstrates how this is done. First, it must be recognized that the map is not actually stored as a single frame picture of a map. Instead, it is stored as a series of feature separates each of which is a frame containing only one feature, e.g., roads, railroads, contour lines, rivers, and so forth. Each such separate is like a black and white acetate overlay. If a pixel intensity is

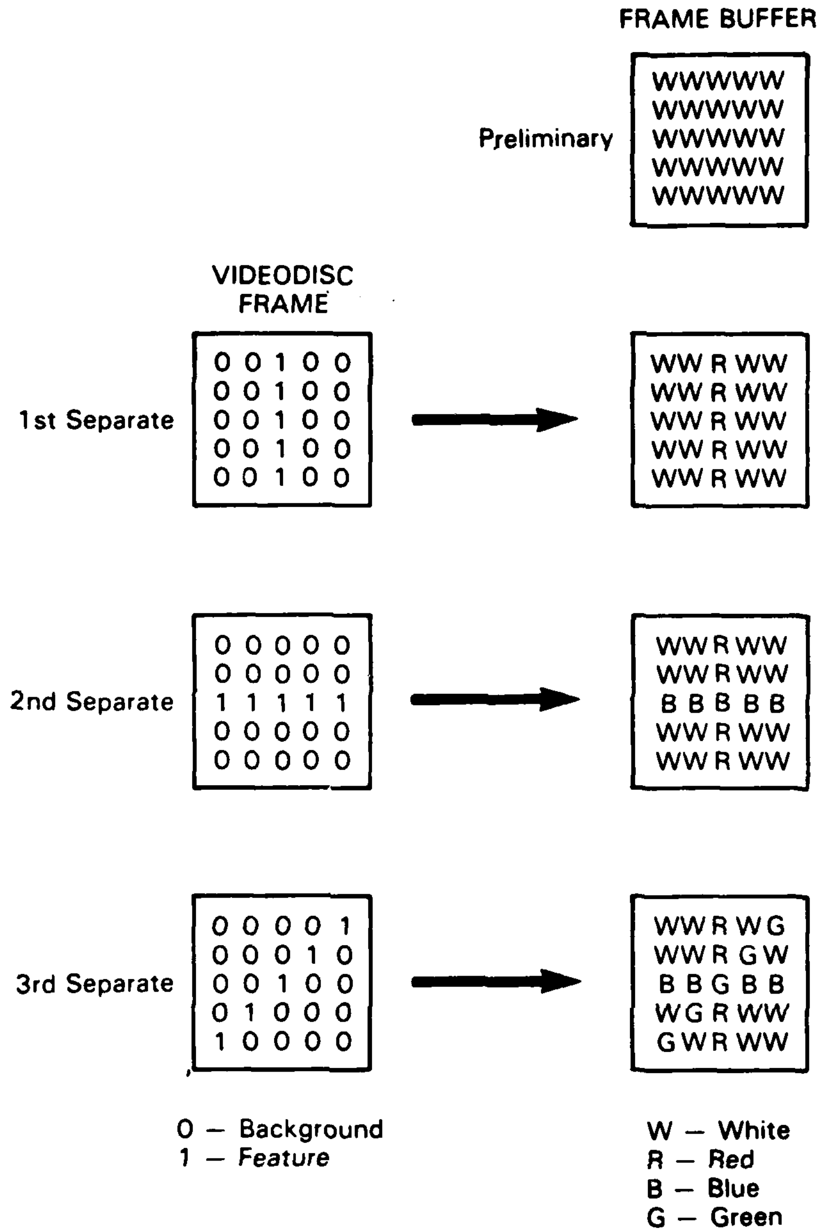

Figure 2-4
EXAMPLE OF OVERLAY TECHNIQUE

black, it contains no evidence of the feature. If white, some portion of the feature is present at that pixel. Figure 2-4 depicts three such feature separates: a vertical, horizontal, and diagonal line, with 0 representing background pixels and 1 representing the feature, respectively.

Next, it must be recognized that the features for any map are stored consecutively on the videodisc in order of their overlay priority. Consecutive storage assures that the separates can be accessed rapidly once the first frame is located by a random access search. Ordering in terms of overlay priority assures that the features are combined properly so that features such as printing or roads obscure others such as contour lines, rather than the reverse. In Figure 2-4 the diagonal line is deemed more critical than the other features, therefore it is last in the sequence.

Finally, it is the image processor contained within the mapping system that provides the ability to combine the feature separates. As each separate is accessed from the videodisc and placed into the frame buffer, it is first read through the image processor. This device has the capacity to determine whether a pixel contains or does not contain a portion of the feature. If no evidence of the feature is present then the frame buffer is left undisturbed. Otherwise the pixel in the frame buffer is changed to a color value unique to the feature. As the successive features are read, the frame buffer accumulates an image with each additional feature overlaying its predecessors. In Figure 2-4, the end result is a green diagonal line, on a blue horizontal line, on a red vertical line, on a white background. To control the set of features being displayed it is sufficient simply to skip the unwanted features when building the image.

Informal observations of this clutter control technique raise a number of issues concerning its value. First, there is an issue of the adequacy of the source material to provide this capability. Section 2.3 discusses this issue in greater detail, but at present it is sufficient to recognize that the feature separates are video images of the acetate reproduction materials used by the Defense Mapping Agency (DMA) to construct maps. Use of these materials leads to the following problems:

1. In many cases DMA's materials are color separates, not feature separates. Therefore, all features of the same color on the printed map, e.g., black printing, black railroads, black road outlines, are combined within a single separate.

2. DMA often uses line screening or dot screening techniques to achieve shading within their reproduction materials. These are periodic arrays of fine dots or lines. When combined with the raster array of video they can introduce moire patterns.

3. DMA's rules for combining separates during printing are not always the simple overlay rules assumed by the feature control algorithm. In some cases, special masks are applied to protect already printed regions prior to inclusion of a new separate. In other cases, the result seems to become a blend of the separates.

Based on DDI's experience, only one map overcame most of these difficulties and it was not a standard DMA map. The more traditional 1:250,000 and 1:50,000 scale maps encountered the difficulties cited above. In principle, it should be possible to obtain maps that conform more closely to our requirements, since DMA begins with feature separates not color separates. These are likely to be more difficult to obtain.

The value of feature control is also challenged by the manner in which the technique degrades the image quality. In particular, the treatment of pixel values as binary introduces difficulties. Aliasing or staircasing of contrast boundaries is a common result. More subtle, however, are the problems that arise due to the difficulty of identifying a suitable intensity threshold for the video image. Each pixel of the video image can actually assume a variety of values. They are rendered binary by assuming that all values below a certain threshold value are background and all values above the threshold are feature.

Figure 2-5 illustrates the difficulties. If all pixels containing feature information are combined to form a histogram of intensities and all pixels containing background information are combined to form a histogram, than each can be thought to reflect a roughly bell-shaped distribution with distinctly different means. The ideal situation is to set the threshold roughly midway between the means. When it is set too low, too much background is interpreted as feature, with the result that lower priority features are inappropriately obscured. When the threshold is too high, too much of the feature is missed with the result being a faded or overly thin feature. This entire problem is rendered especially difficult by the fact that the proper threshold varies from map to map and within a map due to lighting variations during filming.

The solution to these problems seems to lie in gaining increased control over the data capture process (see Section 2.3). In particular, it is necessary to create extreme differences in intensity between background and feature

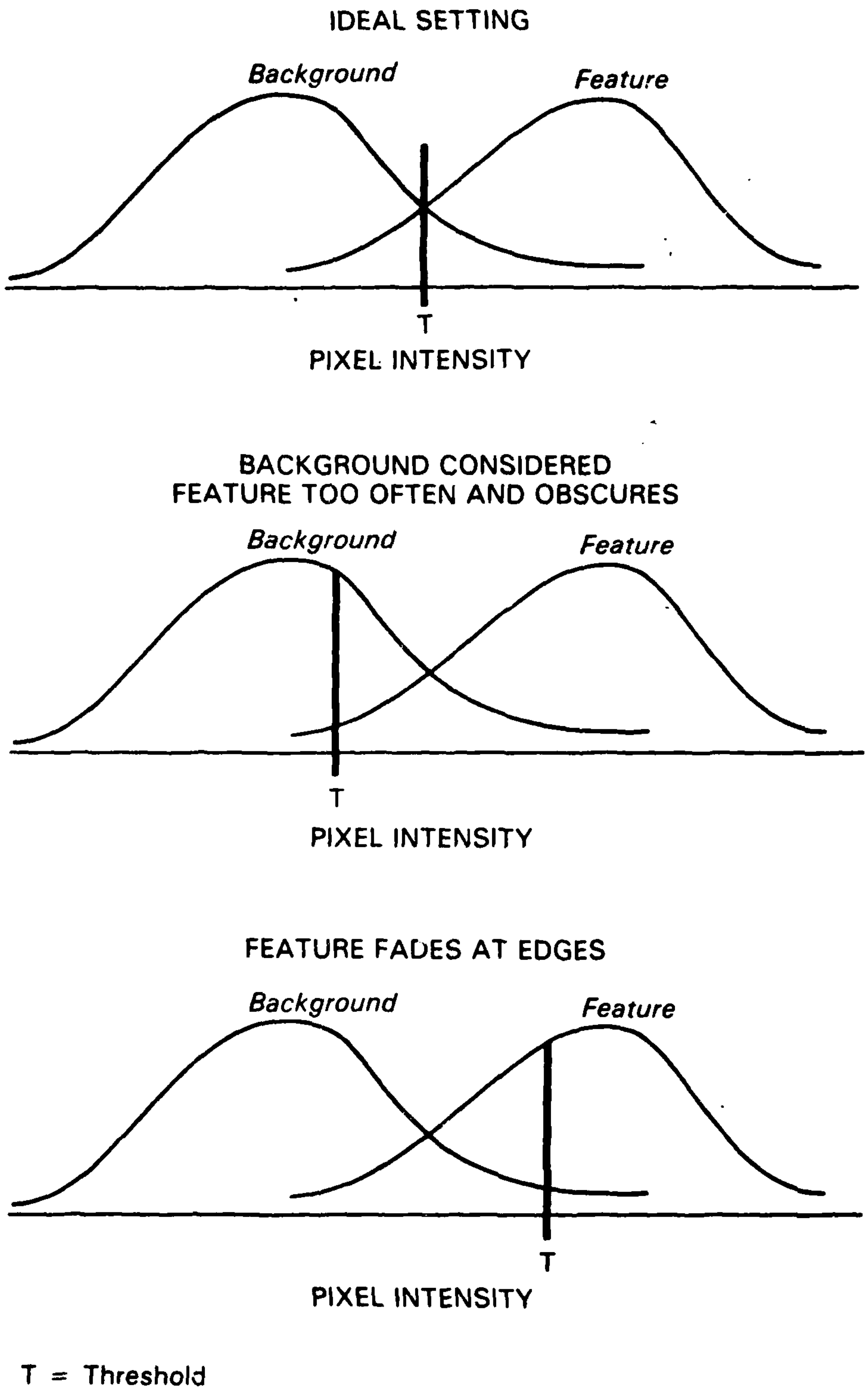

Figure 2-5

SETTING A PROPER INTENSITY THRESHOLD

pixels. Also it is desirable to allow more than just one intensity, given that the feature is present. This latter capability reduces the aliasing problem.

A third issue raised by the clutter control technique is its hidden value as a means of providing color while using a frame buffer. Although video signals carry color information, this information can be difficult to digitize in a frame buffer. To combine color video and frame buffers it is necessary to use a decoder that translates the video signal into three video signals representing red, green, blue intensities. These three signals can then be digitized and placed into three frame buffers for display. At a minimum, this technique requires four bits per pixel for each color to provide an adequate display and it does not provide feature control. In contrast to this 12 bit per pixel approach the use of as many as 32 features requires only 5 bits per pixel, if the present feature control technique is used.

The ultimate question seems to be whether feature control can justify the difficulties it introduces. On the positive side it is a useful capability whose value is not yet determined. Also, assuming a frame buffer will be used, it simplifies the display of color. On the negative side, however, feature control degrades the image quality and is expensive, since it requires both a frame buffer and an image processor. These considerations suggest that the capability is not yet fully developed and requires additional research.

2.1.4 Access to photography - The final computer atlas capability incorporated into the Geographical Display System is the ability to access photography. This capability is quite simple to provide. Aerial and other photography associated with specific facilities or geographic features are stored on the videodisc. Then, by touching the displayed map at a location for which photography is available (this is indicated by a symbol) the requested photography is sought and eventually replaces the map as the display. Return to the map is accomplished by once again touching the CRT.

This capability to access photography is simple to implement and definitely worthwhile. In essence, it is an adaptation of Negroponte's (1979) notion of spatial data management with the map providing the organizing framework for storage of photographs. Admittedly, the photographs must be prestored, unless additional equipment is provided, e.g., a magnetic medium for storage or a videodisc with write-once capability. Nevertheless, access to photography is a valuable addition to the computer atlas.

2.2 Interfacing the Computer Atlas with a Data Base

As described so far, the computer atlas is simply a filing system for pictures. Admittedly, the geographical relationships among the pictures are being fully exploited in an effort to provide a more natural means of access. Nevertheless, the data are the pictures.

A computer atlas provides, however, a capability not available in paper atlases. This is the capacity to interface the maps with a data base and use them as the medium for presenting and understanding the implications of the data base. To appreciate this ability three aspects of the geographical display system must be understood; namely:

(1) computer overlays
(2) color tables
(3) the touchscreen

2.2.1 Computer overlays - One of the primary features of a frame buffer is the ability of its host computer to alter its contents. Thus, there are two paths by which information can be placed in the frame buffer: one via the frame grabber (A/D converter) permits video images to enter, the other via the computer permits portions of an image to be changed.

The computer's capability to alter the pixel intensities stored in the frame buffer is quite general. It can draw a line, fill in a region, or if necessary fill the entire frame buffer. In the case of the target nomination aid's data base, the computer must write symbols into the frame buffer wherever a target is located within the currently displayed region. Each symbol is approximately 16 x 16 pixels and presents an icon associated with the type of target present at the location. Since the computer writes its information after the background map is assembled, the target symbols are essentially the final and highest priority overlay.

This technique for overlaying symbols upon the maps is simple and quite effective. Indeed, the only problem arises when the scale of the display changes. For very small scale maps containing numerous targets, and therefore symbols, the map becomes virtually obliterated when the 16 x 16 symbols are overlayed. The current solution to this difficulty is to use a small dot to represent the presence of a target when displaying small scale maps. Unfortunately, this sacrifices the ability to indicate target type with the symbolic icons. An alternative approach is to record a hierarchy of importance for targets which would permit the computer to filter out the less important targets when displaying small scale maps. In military applications this seems especially feasible, since the command hierarchy provides an appropriate single element that can represent its subordinate elements when they cannot be displayed. For example, Division command posts can represent their subordinate units on small scale maps.

2.2.2 Color tables - An important feature of most frame buffers is the color table. This is a lookup-table that is used as each pixel is read out of the frame buffer and routed to the display. The color tables permit the user to translate pixel values almost arbitrarily prior to display. Thus, the intensity values stored in the frame buffer need not correspond to colors and can instead correspond to a more meaningful or useful dimension.

Figure 2-6 depicts the operation of a color table. In this case each pixel within the frame buffer can assume one of 256 values (8 bits). Therefore, the color table consists of 256 registers, each capable of holding an 8 bit value. The color table has been loaded by the computer to map odd intensities into white and even intensities into black. Thus, when a pixel value such as 17 is read from the frame buffer it is replaced by the value 255 prior to display.

This color table aspect of frame buffers is more than just an interesting detail. It allows one to make a distinction between the information stored in the frame buffer and the color in which that information is eventually displayed. Because of this distinction, the system designer quickly gains the capability to change the color of objects without modifying the frame buffer.

In the geographical display system the color tables are used extensively. One use is to provide color to the feature separates discussed in Section 2.1.3. In that case the separates were assigned intensity values based on their order of overlay more than their color. The color tables were then set to provide the proper color.

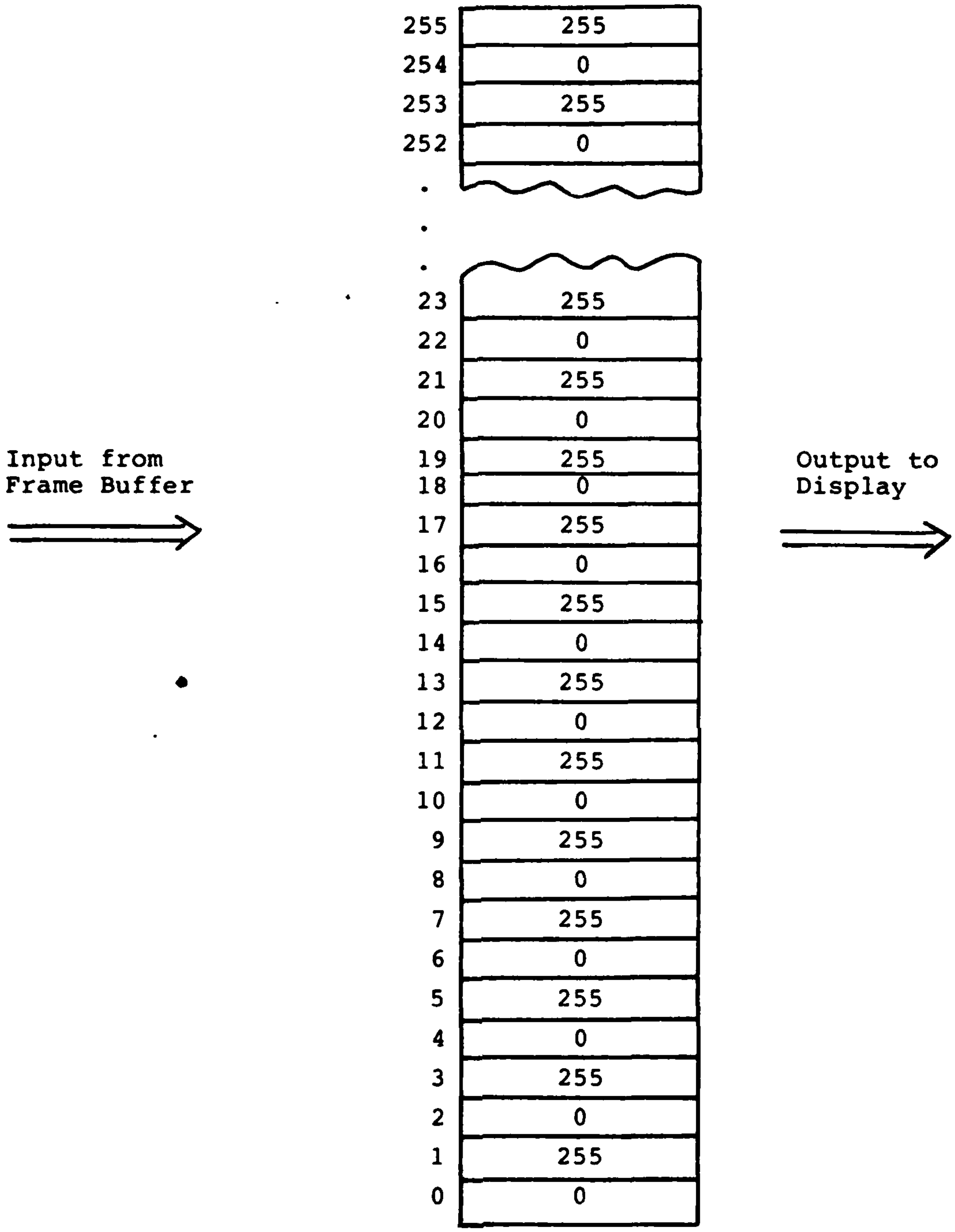

Figure 2-6

COLOR TABLE TO SET ODD INTENSITIES
WHITE AND EVEN INTENSITIES BLACK

A more interesting use of the color tables is related to the symbolic icons used to represent the targets. When these 16 x 16 pixel symbols are written into the frame buffer each pixel takes on one of two values. The foreground, which appears black on all symbols, is written into a common value regardless of target. The background, however, is written into a unique value representing the particular target. The reason for this scheme is that it permits the backgrounds to change color rapidly. In particular, the backgrounds can be colored according to data attributes associated with targets simply by changing the color tables. Thus, a rapid ability to display target data is provided.

2.2.3 The touchscreen interface - To facilitate interaction between the user and the maps, a touchscreen was incorporated into the geographical display system. This device is placed on the face of the CRT and sends a signal to the computer when touched. The signal indicates the location at which a touch occurs and can, therefore, be used to identify geographic locations of interest to the user.

At present, the touchscreen is being used to enter targets, request photographs (see Section 2.1.4), request a data printout for a designated target, and to draw circles upon the maps. Once calibrated, the touchscreen has proven quite effective for most of these control operations. It is doubtful, however, that it could be used for mensuration purposes.

2.3 Map Data Capture

An often overlooked, yet critical, aspect of a geographic display system is the process by which the maps are captured and thereby made available to the computer. In the case of a videodisc system, this process deserves especially close attention both because it is complicated and because alternative approaches may be preferable.

The procedure for placing the maps on the videodisc involves three steps:

(1) acquisition of the source material
(2) filming of the source
(3) videodisc mastering

Each step corresponds to a different medium for storing graphic information. Thus, the overall process is rather complicated and is influenced by the incompatability of the various media.

The original material for the present project was copies of the pin-registered acetate reproduction material used by DMA to print maps. The four basic maps were:

(1) 1:700,000 with 12 separates
(2) 1:250,000 with 11 separates
(3) 1:50,000 with 5 separates
(4) 1:12,500 with 5 separates

Since the 1:700,000 map was too large to fit on the animation stand that would film the separates, it was reduced by DMA to a 1:1,000,000 scale before delivery to DDI.

As discussed earlier (see Section 2.1.3), these maps were inadequate in a number of respects. First, they were not truly feature separates, but color separates. Their reason for being separate is that each is used in turn during the lithographic process to add each new color to the map. Thus, in many cases all features of the same color are combined within one separate.

A second inadequacy is that these materials often use dot or line screening techniques to introduce shading. Although these regularly spaced lines or dots work quite effectively when producing paper maps, they tend to introduce moire patterns when eventually encoded as video images.

Finally, the reproduction materials are not truly overlays. For example, the road separate will contain gaps wherever some other separate, such as rivers or railroads, intersects it. For the geographical display system it is preferable for each separate to be complete. Intersection and contention among separates can be resolved in the computer. This is not true of the lithographic process. Indeed, in some cases a mask must be used before combining certain colors in order to protect the previously printed areas. These complicated combination rules for separates lead to peculiar results when ignored by the geographical display system.

The next step of the map data capture process is to send the source material to be filmed. This filming requires the use of a computer-driven animation stand, preferably with a light table for back lighting and pin registers. The animation stand is a table that can be moved in precise steps underneath a camera that can be raised and lowered. A computer must be in control of both directions of lateral movement, the camera shutter, and the film advance. The light table with backlighting assures sharp contrast and relatively uniform lighting. The pin registration assures that each separate can be placed in the same position on the table.

For any given set of separates the procedure is as follows:

(1) Place a separate on the table.
(2) Set the camera to the upper left corner.

(3) Shoot one frame and advance the film n-1 frames (n = number of separates).

(4) Move 1/3 of a frame to the right.

(5) Repeat Steps 3 and 4 until the rightmost edge of the separate is encountered.

(6) Set the camera 1/3 of a frame down from the previous row and at the far left.

(7) Repeat Steps 3 through 6 until the lower edge of the separate is encountered.

(8) Load a new separate and roll the film back to the frame following the first shot for the previous separate.

(9) Return to Step 2.

In essence, this procedure scans across the separates in a left to right pattern that moves down on each pass, ensures the proper amount of overlap, and interlaces the separates so that each map segments' separates are placed sequentially on the film.

For the present application, six series of maps were produced. In all cases the map segments were filmed on 35mm black and white film in academy size (4 to 3 aspect ratio). Table 2-1 provides the take sizes for each series. Only the 3.2" x 2.4" and 4" x 3" take sizes were sufficient to reproduce lettering on the separates. For the larger take sizes only separates representing regions, e.g., oceans or gradient tints, rather than lines, were usable.

The filming is actually one of the least error prone aspects of the data capture process despite its seeming complexity. The separates aligned well and the overlap was uniform despite the fact that the light table would become hot, thereby introducing the possibility of an expansion of the separates. The primary source of error at this point seems to

Series	Map	Size of Take
1	1:1,000,000	25.6" x 19.2"
2	1:1,000,000	12.8" x 9.6"
3	1:1,000,000	3.2" x 2.4"
4	1:250,000	3.2" x 2.4"
5	1:50,000	4" x 3"
6	1:12,500	4" x 3"

Table 2-1

TAKE SIZES FOR SIX SERIES OF MAPS

have been non-uniform lighting effects both within and between map frames. As explained in Section 2.1.3, this non-uniform lighting can introduce problems for the technique by which separates are combined.

The final step of the map data capture process is to master a videodisc. For this step, the film is sent to Disco-Vision Associates or some other mastering facility with the stipulation that it contains still frames and can be played at 30 frames per second. This stipulation is important, since it will otherwise be assumed that the film contains motion pictures and must be played at 24 frames per second. The problem which arises is that video displays at 30 frames per second; therefore, an extra copy of every fourth film frame must be introduced to ensure proper timing. At the videodisc mastering facility the film is first transferred to videotape. Then this videotape is used to create the stamp mold, which is used to create copies of the videodisc.

The problems introduced by videodisc mastering are of two types. The first is that although mastering requires less than a week, one can wait four or five months for the process to be completed. This is due to the fact that there are so few mastering facilities and it is difficult to get into their queue. More mastering facilities are being built and this may alleviate this difficulty.

The second problem is that accurate random access at all times cannot be assured. By and large, the random access capability is quite accurate, but occasionally it leads to an erroneous frame. When scrolling, this can lead to a suddenly meaningless display.

3.0 TARGET NOMINATION ANALYSIS

The target nomination analysis takes information about user preferences and mission characteristics, along with information about target characteristics, and nominates a set of targets which offer the greatest expected value of damage, given the number of sorties available. The user supplies information about the types of targets which are most valuable by selecting a mission objective and specifying how many sorties are to be allocated to this objective. In addition, the user indicates the approximate geographic area of major concern, and the time when the maximum effect of the mission is to be felt. This information is combined with target information stored in a target data base to obtain a measure of the value of a target to the particular mission at hand. Targets are nominated on the basis of the value obtained, per sortie allocated. The user has the capability to modify model parameters in certain, specified ways.

The intended application of this analysis is in nomination of targets for planned missions. Its chief use would be in offensive counter air and deep interdiction missions, although it has applicability in other areas, as well. The analysis is not an aid for mission planning, and hence, does not consider factors such as specific aircraft locations, mission suitability of aircraft, available weapons and fusing, and so forth. In addition, the aid is not concerned with the problem of validating the location and characteristics of potential targets from data generated by a variety of electronic sensors, photographs, and reports. The chief goal of the analysis is to provide a framework for integrating objective data about target characteristics with critical information about the preferences of the commander, so that potential targets may be evaluated in a manner that is consistent with mission objectives and the overall battle plan.

The analysis is designed to produce a complete nomination list from minimal user input. Thus, to be successful, the model must embody the general knowledge of experts on tactical Air Force targeting. A decision-analytic model is used as a framework to represent expert knowledge. The analysis integrates expert knowledge with specific information about objectives and preferences to produce a first attempt or "straw man" target nomination list. It would be unusual if the initial model had anticipated all the details of the specific situations in which the model is used. Consequently, it is necessary that the user have some method for adjusting the model to reflect the current situation more accurately. The procedures which have been developed for adjustment of the model in response to user feedback are given in Section 4.0.

This section describes the decision-analytic model of target value and the resource allocation model used to develop the initial target nomination list. The analysis is first described from a user perspective; then specific methods, assessments, and procedures will be described.

3.1 User Perspective

The interaction between the user and the decision aid begins with a short procedure in which the aid interrogates the user about the mission objective, the number of sorties to be allocated to the objective, and the preferences concerning the time and geographic area in which an effect is desired. The analysis then combines the information obtained from the user with target information stored in the target data base to develop a target nomination list. The user has several options for examining data and calculated results. The following subsections describe the assessment procedures, the target data base, and the options for examining results, respectively.

3.1.1 Assessment procedure - The four variables which are assessed from the user are the mission objectives, the number of sorties allocated to the mission, the geographic preference, and the time preference.

3.1.1.1 Mission objective - Mission objectives are selected from a set of seven possible objectives. These objectives should not be confused with the types of targeting missions (e.g., offensive counter air, close air support, deep interdiction) although objectives are not entirely unrelated to mission. Rather, the objectives refer to the type of enemy system that should be destroyed, such as the command and control system, the force projection system, the communication system, and so forth. The seven objectives, currently being considered, and their definitions are given below.

1. Blunt the Enemy Offensive. This objective is concerned with destroying those targets that are directly related to the enemy's ground battle. Targets that are particularly relevant to this objective are those that are primarily concerned with projecting force in the ground battle, e.g., infantry regiments, armored regiments, artillary regiments, division and corps command posts, and so forth.

2. Counter Air. This objective is concerned with targets relevant to the enemy's air battle. Relevant targets include air fields; forward air control; petroleum, oil, and lubricant (POL) storage sites for air; surface-to-air missile (SAM) sites; and air defense headquarters.

3. Counter Sea. This objective is concerned with targets relevant to the enemy's sea battle. Although the sea battle plays a very minor role in Korea, this

objective was included for completeness. Relevant targets include naval bases; ports; submarine dispersal stations; naval command, control, and communications (C3); naval supply, and naval training.

4. Interdict the Battlefield. The goal of this objective is to isolate the battlefield from necessary supplies. This objective is concerned with the disruption of lines of communications (LOC), such as bridges, road intersections, railroad marshalling yards, railroad bridges, etc.

5. Degrade Air Defense. This objective is concerned with targets that provide the enemy the capability to defend itself against our air missions. The targets that are most relevant to this mission are SAM sites, air defense headquarters, air fields, air POL storage sites, and so forth.

6. Degrade Command, Control, Communications, and Intelligence (C^3I). The goal of this objective is to disrupt the enemy's C^3I network. Relevant targets include command headquarters, corps and division command posts, and communications links such as radio, television, and microwave.

7. Attrit Reserves. The goal of this objective is to destroy the enemy's reserves, as well as the capability to produce additional reserves. Relevant targets include POL and munitions storage sites, power generation stations, and industry.

These objectives represent a first attempt to classify objectives in a way that distinguishes important from unimportant targets. The success of the overall analysis does not depend on the exact way in which the objectives were defined. In fact, it is anticipated that different commanders will have different and idiosyncratic definitions of mission objectives. However, success does depend on the ability to classify objectives in a way that defines clusters of relevant targets, and that specifies an order of importance on the targets in each cluster.

3.1.1.2 Number of sorties - The user selects a single mission objective. To this objective he assigns a number of sorties. In situations in which there are several objectives, the user may perform a separate analysis for each objective, assigning the appropriate number of sorties to each analysis.

Currently, target nomination is made with little, if any, concern for the number of sorties required to achieve the desired level of damage for any particular target. The sortie requirements are included in the target selection and mission planning processes. Since this aid is chiefly for the target nomination process, details of the types of aircraft and their suitability for the desired mission were not included in the model. The number of sorties, therefore, represents a standardized number, reflecting the current composition of the friendly force.

3.1.1.3 Geographic preference - The geographical preference indicates where the most important targets are located. The user specifies the geographic preference by drawing with his finger a closed curve around the area of interest on the map display. The touch screen will read the set of points indicated by the user, and the points will be echoed to the user on the map display.

The analysis will assume that the primary interest of the user is in targets in the center of the closed curve (which is approximated by an ellipse). Geographical preference will decrease as the distance from the center of the ellipse increases. There will still be some positive preference for targets outside the area indicated by the user, and some may be selected if they are particularly valuable for some other reason.

3.1.1.4 Time preference - A traditional analysis of time preference assumes that the optimal time for an effect is immediately, and that the value of an effect decreases with the delay of its onset. However, there may be cases in which a delay in the effect is desirable; for example, when it is desired that the effect coincide with other plans, such as plans for a ground offensive. The current analysis generalizes traditional methods to include cases in which delay is desirable.

In order to perform the current analysis the user must specify a window of acceptable delays for the effect of a mission. The user specifies two times: the time at which the maximum effect is desired, and the time after which an effect would be too late. The model infers a time preference function from these judgments. The time preference function has its maximum at the desired time of maximum effect, and is sufficiently small at the time judged to be too late. As is the case with geographic preferences, a target may be nominated even if the effect does not fall in the range specified by the user, if it is especially valuable for other reasons.

3.1.2 Target data base - One of the major benefits of decision analysis is its ability to integrate important subjective factors in a rigorous, quantitative manner to find optimal solutions to problems. The fact that subjective factors are important in targeting analysis was stressed in the Introduction to Air Force Targeting (AFP 200-17, 1978). However, subjective factors are seldom considered in a quantitative analysis of targeting problems. Since decision analysis offers the possibility of integrating these factors into the analysis, emphasis was placed on subjective factors in this project.

The targeting analysis currently operates with a data base consisting of twelve factors recorded for 363 targets. The data base represents a hypothetical distribution of targets in North Korea. Although it contains only notional values, the data base was designed to be as reasonable as possible so that the functions of the model could be illustrated.

Most of the factors considered are described in AFP 200-17 under the heading "Qualitative Analysis". The methods developed in this project allow these factors to be incorporated in a quantitative analysis. The following twelve factors include one factor which assesses the cost (number of sorties) required by the targets, and eleven factors which are used to determine the value of the targets. The target data base consists of scores for each of these factors for each of the targets.

1. Cost. This factor is an estimate of the number of standardized sorties required to achieve 70 percent damage.

2. Local Defenses. This factor rates the extent of local defenses for a target on a six-point scale from zero to five. A score of 0 indicates no local defenses; a score of 1 indicates defense by small calibre Anti-Aircraft Artillery (AAA) with a visual Fire Control System (FCS). At the other end of the scale, score of 5 indicates defense by multiple SAM sites and large calibre AAA with radar FCS.

3. Currency of Intelligence. This factor measures the currency of intelligence on the target in days.

4. Concealment. This factor measures the extent of camouflage or natural concealment of the target on a three-point scale from one to three. A score of 1 indicates no concealment; the score of 2 indicates camouflage, and a score of 3 indicates complete concealment.

5. Mobility. This factor classifies targets according to whether they are fixed, transportable, or mobile.

6. Hardness. Hardness is indicated by the pounds per square inch (psi) overpressure necessary for the required level of damage.

7. Dispersion. This factor measures the geographic dispersion of the target. An indicator of dispersion is the number of aiming points which must be destroyed to attain the required level of damage.

8. Capacity/Importance. This factor measures the overall size, level of activity, and general importance of the target relative to other targets serving the same function. Thus, in scoring this factor, bridges are compared to bridges, air fields to airfields, and so forth. The factor is scored on a zero to 100 scale.

9. Time to Impact. This factor measures the time it would take for the effect of damage to a target to get to the FEBA. It is currently measured in days.

10. Time to recover. This factor measures the time in days it would take to rebuild or repair the target after it is destroyed.

11. Location. The final two data variables represent the longitude and latitude of the target location.

The user may examine each of these factors, as well as a number of calculated results described in the following sections.

The analysis uses a fairly small data base because it consists of subjective data at a high level of aggregation. The problem of maintenance of this data base was not addressed so that more attention could be placed on analytical features. It is realized, however, that the problem of data base maintenance is both very important and very formidable. The use of subjective variables in some ways makes the problem even more difficult because assessment cannot be easily automated (although there are only a few variables that need to be maintained).

3.1.3 User options - Of the seven main options given to the user, one is concerned entirely with the display of the map background. The other six are concerned with display and analysis of target data. In addition to these options, the user has considerable flexibility for examination of the background map, as described in Section 2.0. The main program options are described below:

1. Select Targets. When the user selects this option, the target value and sortie allocation analyses begin. The user is asked for the mission objective, the number of sorties to allocate to the objective, and the preferences for time and geographic location of the effect. The analysis calculates the expected

value of damage to each target and allocates sorties to targets so that the greatest value is obtained for the number of sorties used. The results are displayed to the user by coloring the targets on the map display according to the value obtained per sortie allocated to them.

2. Modify Selection. This option is selected when the user wants to change the results suggested by the decision analysis to conform better to characteristics of the particular situation. The user indicates the most salient changes in the results of the model, and the aid modifies the model using one of the methods discussed in Section 4.0 to bring the results closer to those suggested by the user.

3. Examine Target Pictures. Selection of this option allows the user to look at pictures of selected targets, which are stored on the videodisc along with the maps. When the user selects this option, the targets are colored to indicate whether pictures are available for them. Then, when the user touches one of the targets on the map display, the picture is displayed on the map display monitor.

4. Sort Targets by Data Values. This option allows the user to examine the values of all targets on any single data or calculated variable. Targets are colored according to their values on the variable of interest. The color scales may be defined and modified using the option, Reset Display Parameters, discussed below.

5. Display Target Data. This option allows the user to examine all data variables for a selected target. The user selects a target to examine by touching it

on the map display. The data are displayed in tabular form on the control terminal monitor.

6. Change Map Features. This option is concerned entirely with the display of the map. Selection of this option allows the user to control which map features are displayed and which are suppressed. For example, the user may display only cultural features, or only terrain features, or may construct a map containing the features of primary interest.

7. Reset Display Parameters. This option is used to change the color coding scheme that is used to display target data. The user may display data on both nominal and cardinal variables. The display of a cardinal variable is accomplished by mixing two end-point colors in proportions that depend on the value of the variable.

3.2 The Decision Prototype Concept

The concept of a decision prototype provides an organizing framework for the decision aid. This approach is consistent with the goal of developing an advisory decision aid. Details of the concept are described below.

3.2.1 The basic concept - Research on decision analysis and the development of decision aids has traditionally used two approaches to the application of decision analysis in computer aids. In the first approach, options, evaluated structure, values, weights, and probabilities are all assessed from the user with the guidance of the computer aid. Such an aid is potentially applicable to all situations for which the decision-analytic method embodied by the aid is appropriate. The time it takes for such an aid to come to a solution depends, of course, on the nature of the problem, but it is in the order of hours for any problem of reasonable size. Examples of decision aids using this approach are OPINT (Allen, Kelly, Phillips, and Stewart, 1976), QVAL and Gentree (Weiss, 1980), MAUD (Humphreys and Wisudha, 1979), and GODDESS (Leal and Pearl, 1976).

The second approach to decision-analytic aids is based on the concept of a decision template (Kelly, Gulick, and Stewart, 1980). A decision template is a general decision-analytic model framework or model schema which applies to a limited class of decision problems. For example, in a multi-attribute utility analysis, a decision template would be a particular hierarchical structure of attributes for evaluating options. The user selects the template that is most appropriate to the decision at hand. The computer aid then elicits options, values, weights and/or probabilities from the user. Using a computer aid based on a decision template is somewhat quicker than a general aid involving structuring as well as parameter assessment, but the time involved is still at least one hour for a problem of moderate complexity. Examples of decision aids using decision templates are RSCREEN (Weiss and Kelly, 1980) and TACVAL (Kibler, Watson, Kelly, and Phelps, 1978). The Friendly Decision Aid (Barclay, Esoda, Cox, Peterson, and Weiss, 1981) and the aid developed by Merkhoffer, Robinson, and Korsan (1979) encompasses both of these decision-aiding concepts.

The decision prototype offers an alternative to the above approaches for situations in which some aspects of the model may be determined in advance of its application. The following four-step procedure is used to develop a decision-prototype aid (see Figure 3-1).

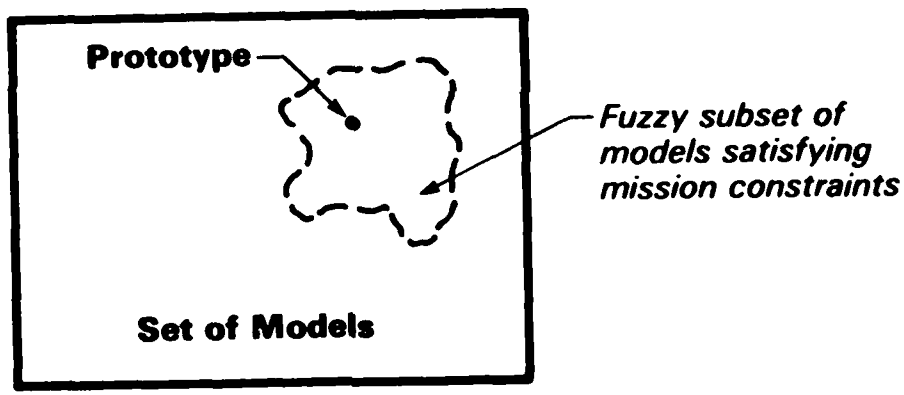

Figure 3-1

PICTORIAL OVERVIEW OF DECISION PROTOTYPE APPROACH

1. The space of models is characterized. Elements of the analysis that are assumed constant are specified. These elements may include the analytical technique, evaluation structures and attributes, parametric forms of combination rules, and ranges on parameter values.

2. Rules for limiting model space are defined. These rules use input from the user to provide a quick way to reduce the scope of the model to a fuzzy subset of the total model space. The purpose of these rules is to make the task of prototype calculation easier.

3. Rules for calculation of prototypes are defined. The prototype represents a complete analysis, and consequently a single point in the space of all models. This point may be determined by the membership function for the fuzzy subset of models specified in the previous step, or by some other method.

4. Rules for adjusting from the prototype to a final model are determined. These rules adjust the prototype to account for additional information supplied by the user.

3.2.2 Application to target nomination - This procedure may be more concrete by illustrating how it is applied to the problem of target nomination. The space of models includes the modeling techniques and the organization of the targets. Two techniques are used to develop the target nomination list. The first is a method that allocates sorties to targets on a value-per-sortie basis. The second is a method for calculating the expected value for allocating sorties to each target.

Targets are organized functionally into components as illustrated in Figure 3-2. Each component is represented by a distinctive target symbol on the map display. For any analysis, each target component receives a non-negative weight which indicates its importance to the mission.

The major user input which is used to limit the space of models is the mission objective. By selecting a mission objective, the user determines a set of constraints on the weights assigned to the various target components (see Figure 3-2). These constraints determine a fuzzy subset of the model space. Other user inputs are used in the calculation of expected values for individual targets.

Two alternate methods have been developed for prototype calculation. The first of these methods picks the prototype model to be that model for which the fuzzy membership function is maximal. The second method bases its choice on a probability measure which is constructed over the elements of restricted model space. The model chosen by this method is the expected value of this probability measure.

Corresponding to the two methods for prototype calculation are two methods for prototype adjustment. The first method constructs a new membership function that integrates user feedback with the original characterization of the restricted model space. The adjusted model is simply the maximum of the new membership function. The second method uses Bayesian adjustment to combine the prior probability distribution over the model space with data from user feedback.

The decision prototype is one way to implement the concept of an advisory decision aid (Patterson, et al. 1981). The model structure embodied by the decision aid gives it the capability to interpret data in a manner that is somewhat

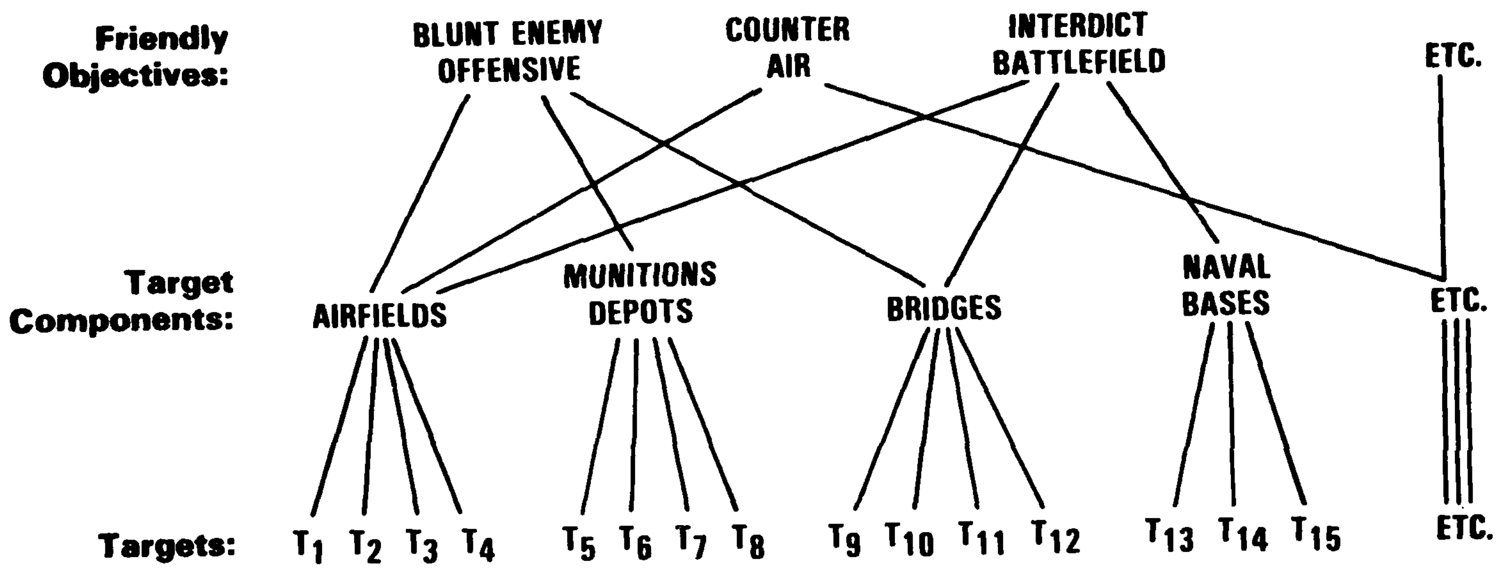

Figure 3-2

FUNCTIONAL ORGANIZATION OF TARGET COMPONENTS different than the methods that are probably used by the user. The aid provides a "straw man" model that may challenge some of the assumptions of the user. However, when the user gives feedback about areas in which the aid is in error, the aid adjusts model values appropriately.

3.2.3 Characteristics of the approach - The decision prototype approach offers several advantages over a decision template. One important advantage is that the aid makes a recommendation, albeit preliminarily, with minimal user input. In the current version of the aid, a nomination list is obtained within five minutes, as compared to at least an hour for the typical aid based on decision templates. A second advantage is that changes are focused on areas where the best guesses of the aid are inaccurate. Use of a decision template forces a user to consider all evaluation attributes. Finally, since the method quickly finds a solution that is close to the correct one, the model will converge efficiently to the correct solution.

The chief disadvantage of this approach is in the time required for development of an adequate model. Use of a decision prototype assumes that the decision aid has some knowledge of the task environment, so that it can evaluate options that were not available at the time the aid was developed. Embodying sufficient knowledge in a computer aid is a much more difficult task than assuming that the knowledge of the user is sufficient to perform the task correctly with help in problem structuring and information aggregation. In addition, the domain of applicability of a decision prototype model is apt to be less than that for a decision template model. Finally, development of a decision prototype contains risk that the model space cannot be sufficiently well-defined so that either the initial model is not accurate, or the adjustment method is not efficient.

As can be seen, the decision prototype concept offers considerable benefits to the user, compared to other methods. However, there are still several areas in which there are risks in developing such a decision aid.

3.3 The Target Evaluation Model

The target evaluation model is concerned with the relative evaluation of targets within a single target component. Four general factors go into the calculation of target value: importance, geographical preference, time preference, and expected damage. Each target receives a score on each of these factors; the overall value of the target is the product of the four scores.

A basic principal in the evaluation of options with uncertain outcomes is the expected value principal. That is, an option should be valued according to the probability-weighted average of the value of its outcomes. Calculation of the expected value of a strike against a target involves the assessment of both a probability distribution and a value function over the extent of damage. Determination of these functions places a considerable burden on assessments. However, it is possible to simplify the assessments if simplifying assumptions are made about the value function. Specifically, if the value of a partial level of damage is a linear function of the extent of damage, then the expected value is the product of the value of complete destruction of the target and the expected damage.

A restatement of the above paragraph in mathematical notation may make the point clearer to the more technically-oriented reader. If D is the extent of damage to a target, then given an allocation of sorties, we may construct two functions V(D) and P(D) which represent the value of damage at level D and the probability of damage at level D or less. The expectation of the value of the damage is given by

$$E[V(D)] = \int_D V(D) \times dP(D)$$

However, if the value of damage is assumed to be proportional to the extent of damage, the value function may be taken out of the integral, and the expectation of the value of the damage is just a constant times the expectation of extent of damage, as shown by

$$E[V(D)] = \int_D k \times D \times dP(D)$$
$$= k \times E(D)$$

In this case, the constant, k, may be interpreted as the value of total destruction of the target.

The value of a target is calculated from importance, geographic preference, and time preference. The overall value is combined with the expected damage according to the equation. Descriptions of the specific calculation procedures follow.

3.3.1 Importance - Importance is currently treated as one of the basic variables in the target data base. The decision was made to treat importance as a basic data variable because it was felt that it would be straightforward to develop uncontroversial indicators of importance that were based on more basic variables. For example, an index of importance for a bridge would combine the number of lanes, the amount of traffic, the span, and the number of alternate crossings. Similarly, an index for an airfield would consider the length and number of runways, whether the airfield was jet-capable, the amount of POL stored there, and the current level of activity. Since the development of indicies of importance was not viewed as an issue for research (although it would involve considerable effort), it was not included in this effort. Consequently, importance is treated as a basic data variable.

Importance is measured on a scale from 0 to 100, where a score of 0 represents a target where no value is obtained from destruction, and the score of 100 is given to the target that has the most value.

3.3.2 Geographic preference - The geographic preference is indicated by the user by drawing a closed curve around the area of interest. However, it does not seem reasonable that this curve should be interpreted literally. That is, geographic preference should be a smooth function of distance from the curve drawn by the user, rather than a discontinuous function which is 1.0 inside the curve and 0.0 outside the curve.

The geographic preference function is shown in this equation.

$$f(x,y) = \exp\left\{ -k/2 \quad \frac{1}{(1-\rho^2)} \right.$$

$$\left. \times \left[\left(\frac{(x-\mu_1)}{\sigma_1}\right)^2 - 2\rho\left(\frac{(x-\mu_1)}{\sigma_1}\right)\left(\frac{(y-\mu_2)}{\sigma_2}\right) + \left(\frac{(y-\mu_2)}{\sigma_2}\right)^2 \right] \right\},$$

where the μ_i, σ_i, ρ, and k are parameters to be estimated from the curve. The function f(x,y), is basically a bivariate normal distribution without the normalization constant. The parameters of the distribution are estimated from the points of the curve drawn by the user. Specifically, the μ_i are estimated by the curve; the σ_i are estimated by the sample standard deviations; and ρ by the sample correlation. The factor, k, is set to adjust f(x,y) so that a point on the curve has about 75 percent of the value as the center of the curve.

It is expected that in practice, the value of k would be tailored to the particular user.

The use of the bivariate normal distribution as a model for geographic preference fits the closed curve drawn by the user with an ellipse. It is expected that this representation would be adequate for a majority of user preferences. However, a number of preferences are not adequately represented by this approximation. Figure 3-3 illustrates three situations in which a curve drawn by a user is represented to a greater or lesser extent by an ellipse.

In Figure 3-3a, the user draws a closed and reasonably regular curve. The approximation of this curve by an ellipse is quite good. In Figure 3-3b, the curve is not convex. The approximation places the area of most concern outside of the actual curve. It seems fair to say that the fit for this case is poor. In Figure 3-3c, the user is trying to indicate an interest for a particular area of the FEBA. In this case the area of maximal interest is a line rather than a single point. The elliptical approximation does not capture this fact about the geographical preference, although an ellipse with enough eccentricity would come close to capturing the preferences. In addition, in the ellipse, preference would decrease symmetrically in front of and behind the FEBA, although this feature is undoubtedly not the case for the actual preference function.

The different mission objectives lend themselves to different representations of the geographic preference function. The current representation is probably most adequate for the objectives, counter air, counter sea, degrade C^3I, and attrit reserves. For the objectives, blunt the enemy offensive and degrade air defense, it may be better to have a preference function with a maximum at a line rather than at a point. For

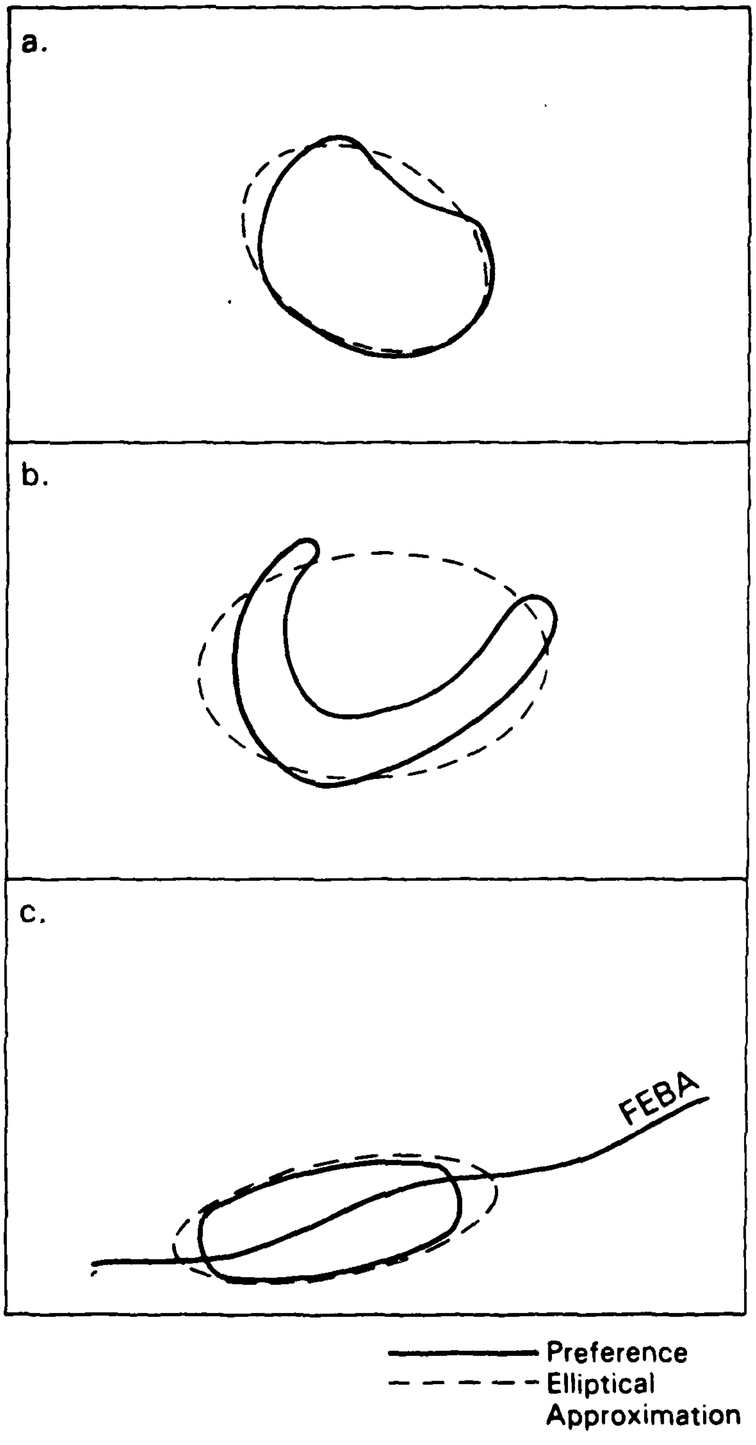

Figure 3-3

**REPRESENTATION OF
GEOGRAPHIC PREFERENCES BY ELLIPSES**

the objective, interdict the battlefield, the method runs into a different problem. It is likely that for this objective, the user would want to indicate the area of the battlefield that is to be isolated rather than the area where the targets are located. The current procedures would not allow this specification, and adapting the analysis to allow it would probably require the use of network analysis methods such as those developed by Kuskey (1981) to determine target value.

3.3.3 Time preference - The representation of time preferences encompasses two concerns: what type of function should be used to represent time preferences, and how should the parameters of the function be estimated. The process of circling an area on a map is a very natural activity; use of this activity made the assessment problem easy in the case of geographic preferences. No such natural activity exists in the case of time preferences. The assessment procedure represents what was felt to be the most natural way to assess the parameters of the time preference function. The form of the function represents a generalization of traditional methods for representing time preferences.

The standard representation of time preferences is a exponential discount function. Under this assumption, the value of a future effect decreases at a constant rate as a function of its delay. Effects with the highest value, naturally, are those that occur immediately, if all other attributes of the effects are held constant. The discount rate parameter of the exponential function expresses the degree to which an immediate effect is important. If it is high, any delay in effect will decrease the worth greatly.

Exponential time discounting was rejected on the grounds that there are instances in which a delayed effect is desired. An example of such an instance is the case in which air strikes are being made in preparation for a future ground offensive. In this case, it seems reasonable that the targeting effort be planned so that targets with a delayed effect are hit early on, and targets in which the effect is immediate are hit immediately prior to the time of the offensive. In this way the onset of the effects of several day's strikes will coincide with the offensive.

Even though it was desired to have a time preference function that does not have its maximum at no delay, another aspect of the time preference function was taken as axiomatic. Specifically, it was assumed that given a constant time of effect onset, a later effect offset was preferred to an earlier one, and given a constant time of effect offset, an earlier effect onset was preferred to a later one. Consequently, the noncentral exponential discount function could be rejected because it does not satisfy these assumptions about the time preference function.

One function providing a reasonable model for time preference is the gamma distribution,

$$g(t) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} t^{\alpha-1} e^{-x/\beta} .$$

The motivation for this distribution comes from the fact that it is the distribution of the sum of independent, identically distributed, random variables. If an immediate effect were desired, then it would be expected that an exponential discount function would accurately represent time preferences. The targets chosen would be those that have a direct and immediate effect on the situation. However, when a delayed effect is desired, targets would be chosen that have an indirect effect on the situation. This effect would probably occur in several stages. If each of these stages is the product of a Poisson process, then the time required for the effect would be distributed according to a gamma distribution. The choice of a gamma distribution seems reasonable in this respect, although this motivation does not give a complete justification.

There are other generalizations of the exponential distribution which may offer reasonable models of time preference. In particular, the Weibull distribution shares many of the properties of the gamma distribution, and is computationally simpler in some cases. The Weibull distribution can represent preferences which are even more extreme than the exponential distribution.

The parameters of a gamma distribution may be estimated using iterative maximum likelihood estimates or, more simply, from the estimates of the mean and variance. Both of these methods are unacceptable because they involve judgments that are not familiar to the user. Instead, the parameters of the gamma distribution were estimated from judgments of the time after which an impact would be too late. The time of maximum input was used as an estimate of the mode of the distribution. The time after which an effect was too late was used to estimate a point two standard deviations above the mean. Chebyshev's inequality ensures that less than 25 percent of the area under the distribution is greater than two standard deviations from the mean. From these two estimates, it is possible to calculate the parameters of the gamma distribution. (For computational reasons, the negative binomial distribution was used as a time preference function in the computer aid. The negative binomial distribution is the discrete analog of the gamma distribution.)

The time preference of a particular target is determined by the time preference function and the time to impact and time to repair from the target data base. Specifically, if t_1 is the time to impact for target i, t_2 is the time to repair, and g(t) is the time preference function; the time preference, TP(i) is given by

$$TP(i) = \int_{t_1}^{t_1 + t_2} g(t)dt.$$

3.3.4 Expected damage - Expected damage is based on six data variables as depicted in Figure 3-4. These variables are used as influencers to determine three factors: the probability of arrival; the conditional probability of finding the target upon arrival; and the expected proportion of damage given that the target is found. The expected damage is the product of the values of these three factors.

Assessments of the conditional probabilities involved in this model should be made from available data and expert judgment. For this model, the assessments represent the well-considered judgment of individuals with extensive military experience, but with no assistance from objective data or analyses. Although it is expected that these figures are generally realistic, it is anticipated that they could be improved substantially, using a more detailed assessment technique.

The probability of arrival measures the likelihood that the plane will arrive safely in the general area of the target. The major influencing data variable describes the local defenses. Assessments were made of the probability of arrival, conditional upon each level of local defenses. The assessed values are shown in Table 3-1. There may be other factors which influence the probability of arrival; an example

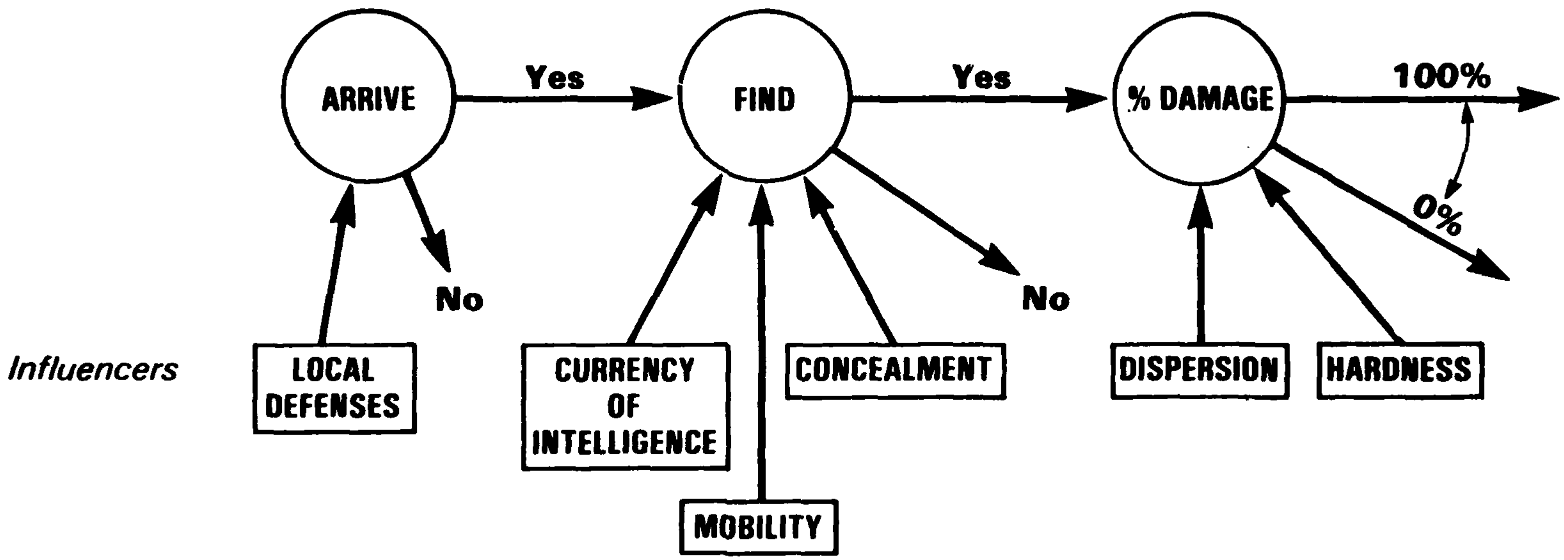

Figure 3-4
CALCULATION OF EXPECTED DAMAGE

LOCAL DEFENSES	PROBABILITY OF ARRIVAL
0	1.0
1	.995
2	.991
3	.987
4	.97
5	.93

Table 3-1

PROBABILITY OF ARRIVAL AS A FUNCTION OF LOCAL DEFENSES of such a variable is the defenses on the route to the target. It would be straightforward to incorporate additional influencing variables into the analysis.

The probability of finding the target measures the likelihood of actually identifying the target, given that arrival in the general area has occurred. This factor is influenced by the currency of intelligence about the target, as well as the mobility and concealment. The assessed probabilities for this factor are shown in Table 3-2.

The expected damage, conditional on the target being found, is dependent on target hardness and on the dispersion (i.e., the number of aiming points). It was judged that at very low levels of hardness, 70 percent damage could be achieved independent of the number of aiming points. For a very hard target (such as one requiring 250 psi for 70 percent destruction), dispersion also would have little effect on expected damage. It is for targets of moderate hardness in which there would be an effect of dispersion. The assessments that were made regarding the influences that hardness and dispersion have on expected damage are shown in Figure 3-5.

The function relating expected damage (E) to hardness (h) and dispersion (n) was approximated by the following power function:

$$E(h,n) = .70 - .5\ (h/250)^{(1\ -\ .756^{n})}$$

The value of this function is shown on Figure 3-5 for selected values of n.

MOBILITY	CURRENCY OF INTELLIGENCE	CONCEALMENT	PROBABILITY OF FINDING
Fixed	1 Day	1 2 3	.99 .98 .90
	5 Days	1 2 3	.99 .98 .90
	30 Days	1 2 3	.99 .98 .90
Transportable	1 Day	1 2 3	.96 .93 .89
	5 Days	1 2 3	.90 .87 .83
	30 Days	1 2 3	.80 .77 .73
Mobile	1 Day	1 2 3	.92 .89 .85
	5 Days	1 2 3	.85 .82 .78
	30 Days	1 2 3	.75 .72 .68

Table 3-2

PROBABILITY OF FINDING TARGET CONDITIONAL UPON ARRIVAL AS A FUNCTION OF MOBILITY, CURRENCY OF INTELLIGENCE, AND CONCEALMENT

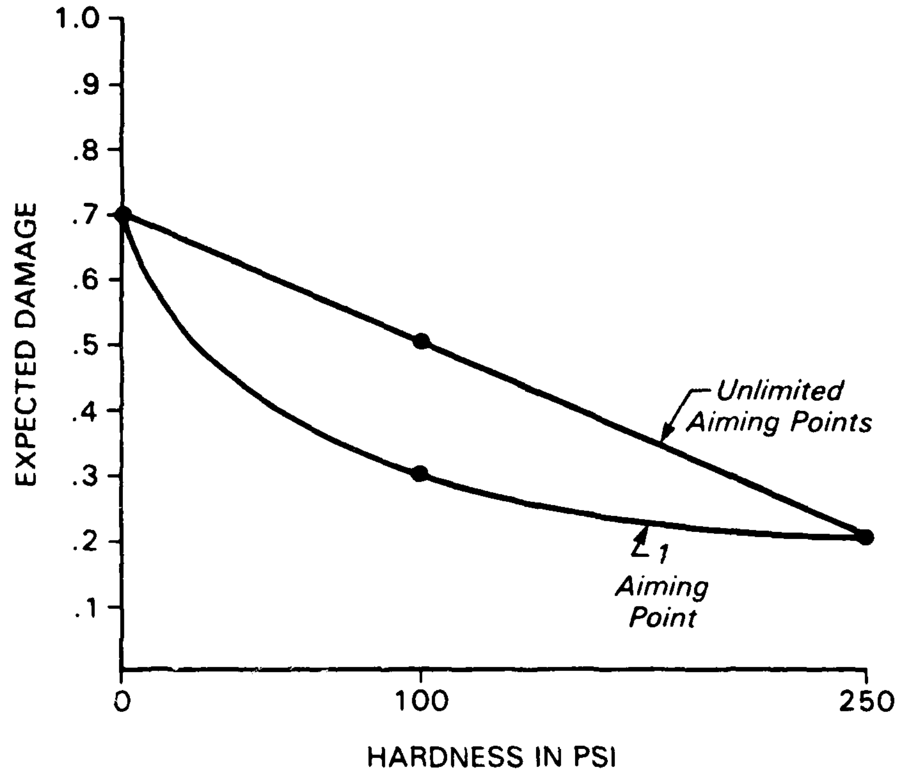

Figure 3-5

EXPECTED DAMAGE AS A FUNCTION OF HARDNESS AND DISPERSION

3.3.5 Target value - The product of importance, geographic preference, time preference, and expected damage is a measure of target value. This measure is applicable only within a given target component. In order to make comparisons between target components, consideration must be made of the weights assigned to the components for the particular mission objective.

3.4 Allocation of Sorties

The goal of the process which allocates sorties to targets is to choose those targets which provide the greatest value to the mission, given the limit on the number of sorties available. The method used to allocate sorties orders the targets in terms of the value to the mission per sorties allocated. Targets are placed on the nomination list in order, until all sorties are allocated. This process guarantees the greatest value will be obtained for the number of sorties allocated (although there may be some unallocated sorties).

The sortie-allocation process occurs in three stages. First, an order of allocation is developed within each relevant target component. Second, corrections are made for certain within-component dependencies. Finally, a between-component allocation is made.

3.4.1 Allocation order within components - The within-component allocation order is obtained by rank-ordering the ratio of the expected value of the target and the sorties required for the target. In future discussions, it will be assumed that the value of the targets for a single component will be normalized to sum to one.

3.4.2 Within-component dependencies - The expected value measure developed above measures target value in isolation of other targets. A variety of interactions between targets may increase or decrease the value of a particular target depending on which other targets are also selected. Many of these interactions are too complex to attempt to model. However, two common interactions seem to be amenable to modeling. These two interactions occur between targets of the same component. Attempts were not made to model interactions between targets of different components.

The two general types of interactions addressed here, "overadditivity" and "underadditivity," are illustrated in Figure 3-6. Underadditivity occurs in situations in which the value of two targets taken together is less than the sum of the values in isolation. A characteristic of these situations is that enemy capability requires all targets to be effective. The reduction when a single target is hit is great; further reductions have much less value. A C^3I network may be a target class in which underadditivity occurs. Destruction of a critical node (or small number of nodes) may have a great effect on the operations of the network. Additional destruction beyond this point has much less value than it would have had if the critical node were not already destroyed.

Overadditivity occurs when the value of two targets taken together is greater than the sum of the values in isolation. These situations are characterized by redundancy in the enemy system. Destruction of one target gets little benefit; it is only by the destruction of several targets that substantial value may be obtained. Air fields may be an example of target components in which overadditivity occurs. If a single air field is destroyed, enemy air strikes may be directed from another one. It is only after several targets are destroyed that the enemy capability for air strikes is severely hampered.

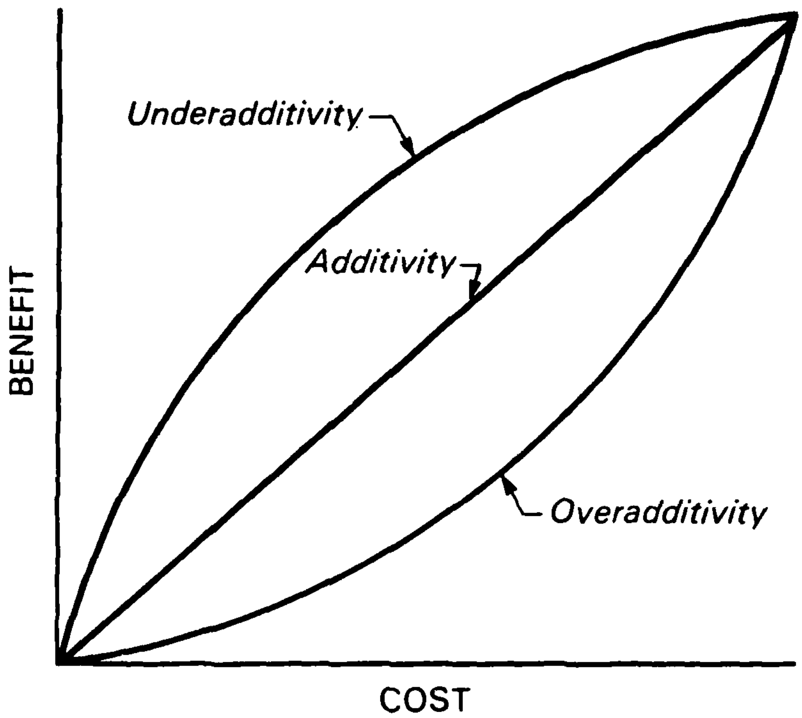

Figure 3-6

ILLUSTRATION OF OVERADDITIVITY AND UNDERADDITIVITY

It is expected that both overadditivity and underadditivity would occur in many target components. However, the model to account for these interactions considers only a single type within each component. Deviations from additivity are modeled by applying a power function to the normalized cumulative value of the targets when taken in the order of allocation. A power of less than one indicates underadditivity, and a power of more than one indicates overadditivity. It should be remembered that the transformation of benefit in isolation by a power function does not affect the allocation order within a component.

3.4.3 Final allocation order - To obtain the final allocation order, the adjusted value for each target of each component is multiplied by the weight for that target component. The weights for each component are shown in Table 3-3 for each mission objective. The result is the ultimate measure of target value to be used to determine the nomination list. The selection procedure is the same as was the case within a component. That is, targets are placed on the nomination list according to the ratio of the value with respect to the mission objective to the number of sorties required.

TARGET COMPONENTS	OBJECTIVES						
	Blunt Enemy Offensive	Counter-Air	Counter-Sea	Interdict Battlefield	Degrade Air Defense	Degrade C^3I	Attrit Reserves
Ministry of Defense	300	200	200	200	300	1000	200
Corps CP	800	0	0	200	100	300	0
DIV CP	400	0	0	100	50	75	0
Artillery Regiment	750	0	0	0	25	25	0
Armored Regiment	1000	0	0	0	25	25	0
Infantry Regiment	500	0	0	150	0	25	0
Road Int.	50	10	0	150	0	0	0
Bridges	125	50	0	1000	0	0	0
Railroad Marshalling Yard	150	50	0	750	0	0	0
Railroad Bridges	125	50	0	700	0	0	0
Army POL	80	0	0	500	10	20	150
Army Munitions	150	0	0	0	150	0	200

Table 3-3

TARGET COMPONENT WEIGHTS FOR MISSION OBJECTIVES

TARGET COMPONENTS	OBJECTIVES						
	Blunt Enemy Offensive	Counter-Air	Counter-Sea	Interdict Battlefield	Degrade Air Defense	Degrade C^3I	Attrit Reserves
Airfields	850	1000	300	300	900	75	25
Air Munition Depot	150	300	175	125	400	0	200
FAC	75	40	0	20	80	10	0
GCI	150	150	150	150	800	100	0
Air POL	75	500	150	100	600	0	150
Aircraft	25	200	100	20	200	0	200
Ports	0	0	900	100	100	75	100
Naval Bases	20	0	1000	100	100	75	25
Naval Supply	0	0	250	50	50	0	400
SAM	75	100	100	100	250	75	0
ADHQ	100	350	100	100	1000	600	0
Navy C^3I	15	0	300	0	100	500	50
Microwave Relay	15	20	15	15	100	300	100
Radio/Television	0	0	0	0	50	300	100

Table 3-3 (Con't.)

TARGET COMPONENT WEIGHTS FOR MISSION OBJECTIVES

TARGET COMPONENTS	OBJECTIVES						
	Blunt Enemy Offensive	Counter-Air	Counter-Sea	Interdict Battlefield	Degrade Air Defense	Degrade C^3I	Attrit Reserves
Satellite RCV	0	0	0	0	0	100	0
Power	15	0	0	15	50	20	750
Industry	10	10	10	100	0	25	1000
Naval Training	0	0	75	0	0	0	100
Sub Dispersal	0	0	900	0	0	0	0

Table 3-3 (Con't.)

TARGET COMPONENT WEIGHTS FOR MISSION OBJECTIVES

4.0 FEEDBACK ADJUSTMENT

The prototype targeting model, because it must be formulated without knowledge of the specific tactical situation, can, at best, capture only general knowledge. It can incorporate information about the relation of target systems to objectives, the rules for evaluating specific targets within a target component, and the approximate importance of each target component given each objective. However, to respond flexibly to the situation-specific goals of the commander and to current target data, the prototype must be effectively adjusted to represent the current situation.

To be effective, the adjustment mechanism must obey the following principles:

- It should allow the user to react to the model in a convenient and natural manner.

- It should adjust the appropriate model parameters to reduce the overall discrepancy between the current model and the user's judgments of the situation.

- The cycle of display-feedback-adjustment should be quick enough to permit several iterations, if the user's responses so indicate.

- The parameter adjustments should attempt to minimize the number of iterations required, and particularly to avoid excessive iteration on a single parameter.

The first two of these principles relate to general principles of the targeting aid's operation, independent of the particular algorithm chosen to perform the adjustment. The third and fourth principles represent a complex set of tradeoffs to be made in the choice of a specific algorithm. In Section 4.1, the general issues are discussed and DDI's approach is defined. The next two sections describe two methodological approaches DDI has studied: an approach based on the theory of fuzzy subsets (Sections 4.2), and an approach based on Bayesian updating of probability (Section 4.3). Finally, Section 4.4 compares the two approaches.

4.1 General Issues in Parameter Adjustment

The overall framework for parameter adjustment in the DDI targeting aid is an iterative cycle in which the current model is displayed, the user reacts to the model by specifying a discrepancy between the model's outputs and his own informed judgment, and then the aid incorporates this feedback by adjusting model parameters to reduce the perceived discrepancy. The aid's model representation at any given time thus reflects a combination of the general expertise embodied in the aid itself; current data about the targets; the user's initially specified objectives, time preferences, and geographic area of interest; and any earlier adjustments due to previous cycles of feedback and adjustment. Disagreement with the model's outputs implies a disagreement with some component of the existing model. Once that conflict is located, a choice must be made: to grant the user's new judgment primacy over earlier information; to accept the new judgment only insofar as it can be reconciled with the original model; or to effect a compromise which may reduce, but not entirely satisfy, the user's objection.

The philosophy behind the targeting aid must be considered in determining how to reconcile the conflict between the current model and user feedback. At one extreme, the information in the original model when applied to the target data and the user's initial preferences may be regarded as the embodiment of a well-tested and comprehensive commander's guidance which the user has authority to modify within a limited range. At the opposite extreme, it is possible to regard the entire model as merely a "straw man," an initial suggestion to be retained, modified, or discarded as the user sees fit.

The approach that was taken is intermediate between these extremes. The recommendations of the general model as applied to the data and the user's initial judgments are regarded as the best application of substantial general expertise, but not as absolute constraints. The user's current knowledge of the situation may identify needs which must be reconciled with the original model, but cannot totally override it. This approach represents both the original model and the user's specific feedback as joint constraints, and automatically selects parameter values to satisfy both as far as possible.

The target symbols on the map display are colored to indicate their position on the nomination list. The user may examine any variable, either basic data or calculated values, to determine the reasons for the result of the analysis. The user indicates his criticism of the order by specifying which targets seem too high or low in the priority order, and where he thinks they should be placed. Because such judgments imply new information about the model parameters, the aid updates its representation of the model to add this new information to its existing model, determines revised parameter values, and then displays the results of the revised model.

The specific definitions of the parameters and the formulas for their updating may differ greatly depending on the particular method chosen. However, the following observations govern all approaches considered during this effort. In all cases, the guidance embodied in the target evaluation rules and in the allocation methods are regarded as fixed constraints that must be satisfied at least approximately by any solution. Target data which reflect current observations can be examined and updated manually if found to be in error, or if superseded by new observations or intelligence; because the data maintenance is essentially an independent function, the aid itself has no automatic capability to make or to suggest changes to this data base.

The user's initial judgments--selection of the objective, the number of sorties available, the geographic area of interest, and the time-of-impact preference--are elicited once. Since these, when combined with the commander's guidance and the target data imply a complete set of priorities, the initial judgments are very important; however, the questions asked are relatively simple and should not change much based on specific judgments about individual targets. (If the user does wish to reassess the initial judgments, the entire order of the targets is likely to change so drastically that such a change would amount to restarting the analysis "from scratch.")

In updating the model parameters based on user feedback, the information which determines the scores of targets within any given target component (e.g., bridges) remains more or less fixed throughout the analysis. Thus, the effect of additional information about specific target priorities must take the form of adjustments to the target component weights. As each new judgment is entered, the aid determines its implications for component weights, adjusts those weights according to a specified algorithm, and calculates a new target priority list based on those new weights.

4.2 An Approach Based on Fuzzy Subsets

4.2.1 General introduction - This approach treats the general guidance from the commander or expert as a set of approximate, general principles to guide the user's actions in accomplishing a given objective. Although this information may be sufficient to imply a specific set of "prototype" weights for the various target components, that prototype is regarded as flexible, as long as certain critical relationships are maintained; more generally, this guidance may take the form of approximate inequalities that produce only a weak constraint on the choice of weights.

Because the target component weights are defined on a ratio scale (i.e., all defining relationships are proportions), the commander's guidance for any given objective will take the form of approximate ratio judgments. For example, if bridges are judged "much more important" than SAM sites in accomplishing the objective "blunt enemy offensive," this might correspond to an inequality such as

$$w_b \geqslant 3w_s ,$$

assuming that w_b is the weight assigned to bridges, w_s is the weight assigned to SAM sites, and "much more important" translates into a ratio of at least 3:1.

In this format, the commander's guidance can be encoded as a set of inequalities which constrain the choice of a target priority order. If many sets of weights can perfectly satisfy all of the inequalities, it is necessary only to find one such set. On the other hand, it is frequently impossible to satisfy all inequalities perfectly; in this case, any given set of weights will satisfy each inequality to some partial degree, which may be quantified in terms of a membership function.

The membership function is a fundamental concept in the theory of fuzzy subsets, which has been applied in the formal representation of approximate reasoning. Consider a relation R which applies to a set of arguments $a_1, a_2, \ldots, a_n$. If a specific n-tuple of values for the a_i satisfies the relation R perfectly, the membership function μ_R, corresponding to R at point $(a_1,a_2,\ldots,a_n)$ will be equal to 1 by definition. Similarly, if a specific n-tuple of values does not satisfy R at all, the membership function will equal 0. If R is defined so that every n-tuple of values for the a_i satisfies R either perfectly or not at all, then the membership function μ_R will equal 1 or 0 everywhere, and the relation R will correspond to a "crisp" (ordinary) subset of the space of n-tuples. An example of such a relation is strict equality, where $\mu_=(x,y)$ equals 1 when $x = y$, and 0 when $x \neq y$.

When the relation R contains a region in which certain n-tuples satisfy R only partially, the corresponding membership function may take on values between 0 and 1. In this case, the set of values satisfying R is said to be a fuzzy subset of the set of all n-tuples; hence R is called a fuzzy relation. For example, if we treat the relation "much greater than" in this manner, one possible definition for the corresponding membership function μ_{MG} might be

$$\mu_{MG}(x,y) = \begin{cases} 1 \text{ if } x \geqslant 3y, \\ 0 \text{ if } x < y, \\ \frac{(x-y)}{2y} \text{ if } y \leqslant x < 3y. \end{cases}$$

In fact, such membership functions in a real problem would be assessed empirically by interacting with the commander or expert, and verified by checking specific examples.

In this context, the commander's guidance generates a set of fuzzy inequalities which serve as rules to guide the choice of target component weights. As additional judgments are made by the user, they too are translated into

AD-A119 076   DECISIONS AND DESIGNS INC MCLEAN VA   F/G 9/5

APPROXIMATE REASONING METHODS FOR DECISION AIDS.(U)

MAY 82 P J STICHA, J F PATTERSON, J J WEISS   F30602-80-C-0156

UNCLASSIFIED   TR-82-6-157   RADC-TR-82-127   NL

2 of 2

END

DATE FILMED

10 82

DTIC

approximate inequalities, and simply added to the set of rules to be followed. The role of the algorithm is therefore to determine what set of weights for the target components achieves the highest joint membership value for all of the inequalities to be satisfied.

4.2.2 Mathematical formulation - The following assumptions underlie the mathematical approach taken in this section:

- Every judgment, including both the expert's guidance and the user's feedback, can be translated into one or more approximate inequalities; furthermore, each of these inequalities relates exactly two component weights. (A judgment of approximate equality can be restated as a pair of approximate inequalities.)

- The membership function corresponding to each comparative relation depends only on the ratio of its arguments; since the weights themselves are defined on a ratio scale, the only meaningful judgments should concern their ratios. More specifically, each relation can be evaluated by comparing the actual ratio of its two arguments to an "ideal" ratio which characterizes that relation. For example, the relation "much greater than" described above would have an ideal ratio of 3:1; any ratio above 3:1 would receive a membership value of 1, while any ratio greater than 1:1 would receive a lesser value between 0 and 1, depending on the actual ratio divided by the ideal ratio.

- The joint membership function (which determines how well a given set of weights satisfies all of the constraints) is determined by that individual inequality which is satisfied the worst. If a given set of weights satisfies all but one of the inequalities, its overall score will be its score on that inequality; to find a set of weights with a higher overall score, the original set of weights should be changed in the direction which better satisfies that critical inequality.

The following standard representation can be used to represent the information needed to solve the weight adjustment problem. Each inequality in the system can be converted to the form, $w_i \geq R_{ij}w_j$ (reversing the order of inequalities as necessary). The R_{ij} values correspond to ideal ratios implied by the judgments. If there are two or more inequalities relating the same pair of arguments (i.e., the same i and j in the same order), only the largest of the R_{ij}'s should be noted; the less restrictive relations can be ignored because they will be satisfied at least as well as the one with the largest ideal value. Naturally, these assessed ratios apply only when $i \neq j$. It is not necessary to have an assessed R_{ij} value for every pair (i,j), as long as the total set of assessments includes at least one chain of assessments $R_{ia}, R_{ab}, R_{bc}, \ldots, R_{yz}, R_{zj}$, for every $i \neq j$. Arbitrarily, R_{ii} will be set equal to 1 for all i, and for any (i,j) pair that is not constrained by any assessed relation, R_{ij} will be assigned a value of 0. The matrix of numerical R_{ij} values completely characterizes the input information for the weight determination problem.

For any given vector of positive-valued weights <u>w</u>, another matrix A can be calculated to represent the actual ratios of the weights to one another, defined as $A_{ij} = w_i/w_j$. Dividing each element of R by the corresponding element of A, we derive a <u>discrepancy matrix</u> D, where

$$D_{ij} = \frac{R_{ij}}{A_{ij}} = \frac{R_{ij}}{w_i/w_j}$$

Whenever $D_{ij} = 1$, the weights w_i and w_j are exactly in the ideal ratio defined by the assessed inequalities. When D_{ij} is less than 1, the ratio w_i/w_j surpasses this ideal ratio, and thus also satisfies the corresponding constraint; the excess represents neither an improvement over equality nor a cause for concern, because beyond the ideal point, the membership function is always equal to 1. Problems manifest themselves when some of the D_{ij} values are greater than 1; these correspond to values of w_i and w_j that do not completely satisfy the given constraints. In fact, the degree to which a given set of weights satisfies the joint set of constraints is a direct function of the highest D_{ij} value in the matrix.

4.2.3 Iterative solution procedure - Given this formulation, one reasonable approach would be to pick an arbitrary set of weights (equal weights, for example), calculate the corresponding A and D matrices, identify the largest discrepancy, and select a new set of weights which will reduce that discrepancy. If D_{ij} is the largest value, it can be reduced by changing w_i and/or w_j in such a way that their ratio increases.

This stepwise process of increasing w_i and decreasing w_j can continue until D_{ij} has been reduced to the point where it equals one of the other elements in the D matrix. Now there will be two elements tied for the highest discrepancy, one of which is the revised D_{ij}. If the other element is D_{ji}, there is no further operation that can reduce D_{ij} without increasing D_{ji} by the same factor (and vice versa); the current ratio of w_i to w_j thus represents the best compromise between two essentially conflicting assessments, represented by R_{ij} and R_{ji}. When this occurs, the ratio of w_i to w_j will be fixed at its current value, and the analysis will proceed by searching for the next largest discrepancy.

On the other hand, if D_{ij} crosses the value of some other element before it reaches D_{ji}, it will be possible to decrease both discrepancies simultaneously. For example, if D_{ij} and D_{pq} are the two largest discrepancies, w_i and w_p should be increased, while w_j and w_q should decrease. This process continues, adding additional elements of D until no combination of changes in weights would permit simultaneous reduction of all discrepancies. When this occurs, a <u>cycle of discrepancy</u> can be identified, consisting of a set of discrepancy values D_{ij}, D_{jk}, D_{kl}, ..., D_{xy}, D_{yi}. (The simplest case of this occurs when D_{ij} and D_{ji} are both among the discrepancies to be reduced, as described above.) Again, the ratios of weights w_i, w_j, ..., w_y must be fixed at their current levels, and only those discrepancies involving other elements of D considered in the effort to locate and minimize the next largest discrepancy. This process continues until the ratios of all weights have been completely determined. It can then be shown that this set of weights (normalized to add to unity) is optimal according to the criterion of minimal discrepancy, as set forth in the problem statement.

This method, while it is conceptually straightforward, requires numerous iterations even when only a few weights must be determined. To make the method practical, a way must be found to identify the critical cycles of discrepancy without stepwise iteration. The next section describes one way to do this.

4.2.4 <u>The cycle identification method</u> - In the method described above, it is possible to reduce the greatest discrepancy by some change in the weights, until a cycle of discrepancy occurs. Therefore, a considerable reduction in time may be achieved if these cycles can be detected using direct analytic methods rather than the iterative numerical technique described above. This section describes some of the mathematical the three elements in the critical cycle. Then, for those inequalities of the form

$$w_2 \geqslant R_{2k}w_k \qquad (k \geqslant 4)$$

would be translated to the form

$$3w_1 \geqslant R_{2k}w_k \quad .$$

Now, if R_{1k} happened to be less than $\frac{1}{3}R_{2k}$, this new constraint would replace the R_{1k} term in the R matrix; on the other hand, if R_{1k} were greater, the other value could be discarded, since the inequality corresponding to R_{2k} would always be satisfied at least as well as the one corresponding to R_{1k}.

All R values corresponding to the elements in the critical cycle are transformed in terms of the chosen surrogate element in a manner comparable to the example above. Then, after noting the (now fixed) proportions which relate the other elements to the surrogate element, those other elements are removed from the R matrix (or, alternatively, their R values are set to zero), and are no longer considered as independent variables, since their value is determined by that of the surrogate element. The result is a new R matrix, transformed and reduced according to the operations described above. This new matrix can be analyzed in the same manner to find the next critical cycle, and so forth, until all ratios have been fixed (i.e., until a critical cycle is found to include all the remaining elements in the R matrix).

To transform the cycle reduction method into a practical algorithm, the only missing item is a reasonably fast method for locating the critical cycle in any R matrix.

properties of these cycles of discrepancy, and one possible way to use these properties in an algorithm.

The first important property of a cycle of discrepancy is a measure invariant under changes in the weights of its components. Regardless of the values of w_i, w_j, w_k, $...w_z$, the product of the discrepancies in the cycle will be equal to $R_{ij} \times R_{jk} \times R_{kl} \times ... \times R_{yz}$, since all of the w terms will cancel. If this cycle happens to be the critical cycle (i.e., if this would be the first cycle reached in the iterative method), the biggest discrepancy is minimized by setting each discrepancy in this cycle equal to the same value Δ (if any discrepancy in the cycle were less than Δ, the invariance of the product would imply that some discrepancy larger than existed as well). Therefore, for all w_i, w_j in the cycle,

$$w_i \cdot \Delta = R_{ij} w_j$$

Note that if this cycle happens to include all of the w_i, these equations, together with the definition of Δ as the geometric mean of the R values in the cycle and the requirement that the weights add to unity, determine the values of the weights uniquely.

Most of the time, of course, the critical cycle will include only some of the weights. When this occurs, the ratios among those weights in the critical cycle are uniquely determined, although their absolute values are not. Nonetheless, it is possible to select (arbitrarily) one of the weights in the critical cycle, and use the determined ratios to substitute for the other cycle elements in the original set of inequalities represented by the R values. For example, suppose the critical cycle implied that w_1, w_2, and w_3 were in the proportion 1:3:4, and that w_1 was chosen as the surrogate for

An exhaustive search would certainly terminate in a finite number of iterations, but the number of possible cycles in an n-by-n matrix grows factorially as n does. Therefore, for even moderately large values of n, some shortcut will be required to make the entire process faster than the iterative numerical method.

Fortunately, significant savings can be achieved by checking partial cycles (i.e., products of the form $R_{ij}R_{jk} \ldots R_{xy}$, where i, j, ... x, and y are all different), and eliminating those that could not possibly be part of the largest cycle. This process involves constructing an N x N x N matrix Q, so that Q_{nij} represents the product of the R values associated with the largest partial cycle of length exactly equal to n (where N is the total number of weights to be determined). The simplification stems from the fact that every partial cycle of length n can be factored into a partial cycle of length n-1, times the value of the R_{ij} corresponding to the larger partial cycle's last component. In other words, for every Q_{nij} there exists a k such that

$$Q_{nij} = Q_{(n-1)ik} \times R_{kj} ,$$

where k is chosen to maximize the value of Q_{nij}.

This observation permits us to construct the matrix Q efficiently, in the following manner: First, set Q_{1ij} equal to R_{ij}. Then construct Q_{2ij}, Q_{3ij}, etc. (in that order), using the above formula. The diagonal element Q_{nii} represents the largest complete cycle of length n that includes the i^{th} weight. The largest Q_{nii} for each n is thus the product of R_{ij}'s for the largest complete cycle of length n.

However, the criterion for selecting the critical cycle is not the product of the R_{ij} values in that cycle, but rather their geometric mean; this is true because each cycle of length n contains the discrepancy ratio as a factor n times. Thus, the proper procedure is to select that n which maximizes the n^{th} root of the largest Q_{nii}. That maximum value will now equal the critical discrepancy ratio Δ.

Now, the elements in the critical cycle can be found by identifying those R_{ij} for which $Q_{nii} = R_{ij}Q_{(n-1)ji} = \Delta^n$. (This step again takes advantage of the fact that the cycle can be factored into an R_{ij} and a partial cycle.)

One possible problem might arise when two or more different cycles have the same critical ratio Δ. However, in that case, a satisfactory result can be obtained by selecting any of the equal cycles, arbitrarily. One way to do this is to select from among the critical cycles the R_{ij} which has the lowest i value, and (if there are two or more j values on that line of the R matrix) the lowest j value. Using this value of j, one would seek the R_{jk} in the critical set with the lowest k value. This process is continued until eventually the original index i is reached. (At this point, the selected R values form exactly one complete cycle.)

To summarize the cycle identification method, these are the essential steps:

1. Represent all inequalities in the form of R_{ij} values, to form the matrix R.

2. Find the critical cycle $R_{ij}R_{jk}\ldots R_{xi}$ by:

    a. Constructing the matrix Q_{nij}, where Q_{nij} represents the largest partial cycle of length n beginning with i and ending with j (for i ≠ j), or the largest complete cycle of length n including i (for i = j).

b. Selecting the largest Q_{nii} for each n.

c. Selecting W to equal the largest value of $Q_{nii}^{(1/n)}$.

d. Identifying those R_{ij} for which $Q_{nii} = R_{ij}Q_{(n-1)ji} = \Delta^n$.

e. Checking for multiple cycles with the same Δ value, and arbitrarily selecting only one such cycle if there is a duplication.

3. Fix the ratios of the elements in the critical cycle, determining the ratios by dividing all R_{ij} in the cycle by Δ.

4. Adjust the R matrix so that the i^{th} row captures all information relating the elements in the cycle to any elements outside the cycle: $R_{ik} = \mathrm{Max}_j \; (w_i R_{jk}/w_j)$, for k outside the cycle, and i = any arbitrary element in the cycle (e.g., the lowest-numbered element).

5. Delete all elements in the cycle except for element i, after recording the (now fixed) ratios which relate the weights for those elements to the weight for element i.

6. Using the reduced R matrix, return to Step 2, to find the next critical cycle; repeat until a weight has been determined for every element in the set.

This algorithm has been implemented on an experimental basis, using an IBM 4331 computer, and has not been integrated with the targeting analysis. The prototype implementation is able to process randomly generated 15 x 15 matrices of R_{ij} values in times consistently under 20 seconds. It is expected that this approach would be considerably slower when integrated with the targeting aid because of the lower power of the host computer, as well as other peculiarities of the system. These facts would not preclude the use of the algorithm in a future implementation.

4.3 An Approach Based on Bayesian Updating

The approach described in the previous section adjusted weights to minimize their discrepancy with a set of fuzzy constraints on the space of possible weights. The approach described in this section is quite different; it defines a probability distribution over the space of weights which it revises using Bayes' Theorem, upon the receipt of user criticism. The Bayesian approach was implemented on the demonstration targeting aid system.

4.3.1 General introduction - The major assumption of the Bayesian approach is that at any point in the analysis importance weights may not be specified exactly, but may vary according to some probability distribution. One way to think about this distribution is to assume that weights depend on many factors which cannot be assessed at any moment. Consequently, the actual weights can never be known exactly. Given knowledge of the mission objective, an expert can estimate what the weights are, but this estimate will not account for situation specifics. Another view of this assumption states that initial weights describe guidance to the targeteer which may be adjusted within limits to account for the specifics of the situation.

Regardless of the interpretation, the approach assumes that there is a probability distribution over the space of weights. At any time in the analysis, the best estimate of the weights is the expected value of the weight distribution. As information is received about the weights, the probability distribution is revised.

New information is received in the form of judgments that a given target is approximately equal in importance to another target. The impact this information has on the probability distribution depends on how likely it is under the probability distribution. Bayes' Theorem specifies the extent to which a probability distribution, $f(w)$, should be revised in light of information, D, stating that:

$$f(w|D) = \frac{l(D|w)\ f(w)}{\int_w l(D|w)\ f(w)\ dw}$$

where $f(w \mid D)$ is the posterior probability distribution of w, and $l(D \mid w)$ is the likelihood of D given the prior distribution, f.

Much of the effort involved in developing a Bayesian updating method is concerned with the determination of an appropriate prior distribution. A conjugate prior distribution has the property that incoming information does not change the basic form of the distribution; only the values of parameters are changed. In this case, the calculations required are made considerably simpler, and can easily be automated. Details of the analysis are described in the following section.

4.3.2 <u>Details of the approach</u> - The specific approach assumes that the component weights are distributed according to a Dirichlet distribution. The Dirichlet distribution is a multidimensional generalization of the Beta distribution (see DeGroot, 1970, for a description of these distributions). A Dirichlet distribution of n variables which are constrained to sum to unity may be described by n parameters, a_1, a_2, ..., a_{n-1}, and A. The expected value of the ith variable, w_i, is a_i/A; w_n may be determined from the values of the other variables. The parameter, A, indicates how spread out the distribution is; high values of A correspond to peaked distributions.

The Dirichlet distribution is an appropriate conjugate prior distribution for the weights. In fact, it may be shown that upon the receipt of data which imply a new set of weights, the new values for a_1, a_2, ..., a_{n-1} are simply the weighted average of the original values and the values implied by the data. The parameter, A, is incremented by a percentage that corresponds to the weight given the data in the revision of the a_i's.

One choice to be made which is crucial to the preparation of the method is the definition of what is a single datum. Two alternatives present themselves:

1. Each user response may represent a single datum.

2. All responses of a user to the initial model results constitute a single datum.

If the first of these alternatives is chosen, the user has potentially unlimited freedom to adjust the weights away from the prior weights. Under the second alternative, this freedom would be limited. This fact would suggest that the second alternative is more appropriate if the prior weights are to be interpreted as the commander's guidance. In the situation in which the prior distribution is interpreted as expert judgment, the choice between these alternatives is much more difficult. For the purposes of the targeting aid demonstration, the second alternative was chosen. One factor that influenced the decision is that the fuzzy-subset approach also limits the extent to which a user may modify the initial weights. The similarity of the two methods in this respect makes unbiased comparison between them easier.

The specific adjustment procedure is described below. As new information is received in the form of judgments, that one given target T_{pi} (the i^{th} target in target component p) is approximately equal in importance to another target (T_{qj}), a set of "adjustment weights" A_t is calculated by the following method:

1. Initially, set all adjustment weights equal to the corresponding prior weights.

2. Adjust the weight A_p for the component corresponding to the first-named target so that $A_p S_{pi} = A_q S_{qj}$ (where S_{tk} represents the score for the k^{th} target in component t).

3. Normalize the adjusted weights to add up to unity.

4. Calculate the updated weights as an additive mixture of the prior weights and the adjusted weights. For example, if the proportional contribution of prior to adjusted weights is M:N, the updated weight U_t on target component t will be defined by the formula

$$U_t = \frac{MW_t + NA_t}{M + N} .$$

5. On subsequent iterations (as the user adds additional judgments of target equality) the prior weights and the ratio M:N remain the same; the only change is the adjusted weights and updated weights. For each iteration after the first, the previous set of adjusted weights will be used in Step 2, instead of the prior weights from Step 1; otherwise, the sequence from Step 2 through Step 4 is repeated each time.

The proportional contribution given to the adjusted weights depends on the extent to which the prior distribution is spread out. Since the uncertainty about the weights is unknowable for the current hypothetical example, the relative contribution of prior weights and adjusted weights was set arbitrarily to give them equal impact on the updated weights. It is expected that the parameter, A, of the prior distribution could be assessed for a more realistic situation, although considerable difficulty would be anticipated.

This approach was straightforward to implement, and fast enough to operate within the context of the targeting aid. It shares the feature of the approach, based on fuzzy subsets, that the user has a limited capability to adjust the initial weights. Specific comparisons between the two methods are given below.

4.4 Comparison of the Approaches

This section contains a comparison between two specific numerical techniques for determining target component weights as a function of initial guidance and the user's feedback. Although these two techniques represent different overall approaches to approximate reasoning (the "fuzzy-sets" and the "Bayesian" approaches), numerous other factors make it impossible to use this comparison to establish a definitive preference for either general approach. Perhaps the most important extraneous factor is the relative complexity of the information stored during the operation of each method: in the Bayesian approach, all information about the weights of the target components is summarized in an N-element vector corresponding to the current best estimate of the weights; in the fuzzy-sets method a richer set of N(N-1) inequality judgments is retained and updated, and this more extensive information can be used at any point to derive the current best-estimate weights. Thus, in interpreting the evaluation of the two specific methods, care must be taken to distinguish the effects of a more complex representation from those of a philosophically different approach.

With the above caveats in mind, this section will compare the two weighting methods in three general areas: the means of representing initial information and user feedback; the behavior of the weight adjustment procedures as the model evolves; and the method's demand on computational resources.

4.4.1 Representation of initial information and user feedback - In the current approach, the initial information (an amalgam of the "commander's guidance" stored in the system, the target data, and the user's initial preferences and objectives) is used to derive a specific initial set of weights for the target components. Since this initial set of weights is the only information retained to summarize the system's "knowledge" prior to the user feedback, both current approaches act equivalently (the fuzzy-sets approach translates the original weights into a consistent set of inequalities which could be used, in turn, to derive those weights). It would, of course, be possible for the fuzzy-sets approach to capture a richer representation of the prior guidance by representing that guidance directly, as a set of inequalities; this would allow the initial guidance to influence not only the original set of weights, but also the directions in which those weights could be modified to be consistent with the original guidance.

In both versions of the present approach, the user indicates disagreement with the current model by selecting one target that appears out of place in the nomination list, and indicating where he thinks it should belong. The fuzzy-set approach interprets this feedback by inferring a set of inequalities among the weights, adding the new inequalities to the retained list of inequalities, and using the augmented set of inequalities to determine a new set of weights. Because these inequalities capture all of the information which can be derived from the user's ordinal judgments, they constitute the most complete possible representation of that feedback.

The Bayesian approach, on the other hand, treats the user's repositioning of the selected target as a statement of equality, and adjusts the weights to give the moved target the same benefit-to-cost ratio as the one which previously occupied the position to which that target was moved. This restriction to judgments of equality could potentially lead to problems when the user wishes to locate a target midway between two widely separated targets. It is probably more natural for the user to think in terms of ordinal inequalities than in terms of specific values for the costs and benefits of targets. In practice, however, the set of targets should be sufficiently dense that the mathematical consequences would be minor.

Looking at the difference between the two approaches in another way, the fuzzy-sets approach represents the user's ordinal judgments directly as inequalities, and retains the specific information in those inequalities in a format that would permit the user to examine the entire set of relevant judgments, in an effort to detect possible errors or obsolete judgments. In the Bayesian approach, information is retained only in the form of the most recent set of updated weights, and cannot be used to retrieve those specific judgments leading to a given conclusion. Thus, the fuzzy-sets approach can fairly be said to retain user feedback information in a more directly useful form.

4.4.2 Behavior of the weight adjustment routines - As additional feedback information is added to the system's representation of the problem, the Bayesian approach uses that information to calculate an updated set of weights, and moderates that updated set by mixing it with the prior weights which represent the original judgment. This operation is quite simple: if the user moves a target of a particular component higher on the nomination list, the weight for that component will be raised somewhat, and all other weights will be lowered proportionally, to maintain a fixed sum of weights.

By contrast, the adjustment of weights in the fuzzy-sets approach is somewhat more complex, as a result of the more detailed representation of information. The fuzzy-sets procedure responds to the new feedback by adding another inequality to its current representation. If this new inequality is approximately consistent with the current representation (i.e., if it does not lead to discrepancies any greater than those which already exist), the weights will not change at all, although the new inequality will be retained as a potential source of inconsistency with future judgments. If, as is more likely, the new inequality does conflict significantly with the current model, an entirely new set of weights will be determined; those target components directly involved in the new judgment will be raised or lowered in accordance with the corresponding inequality, but all other weights that are related by any comparable cycles of discrepancy will also be changed in order to maximize the consistency of the overall system.

Typically, those weights, which are adjusted based on more indirect inferences (i.e., those which were not directly mentioned in the judgment), will change by only a small amount, compared to the changes in those weights directly mentioned.

To summarize the advantages and disadvantages of the two methods, the Bayesian approach is extremely simple and direct, changing one weight at a time; this simplicity is achieved at a cost, however, because the impact of each new judgment is calculated independent of the specific judgments that preceded it. The fuzzy-sets approach permits more flexible changes in the entire set of weights to accommodate the new feedback into the existing model, while maintaining sensitivity to specific previous judgments; although the user would not, in general, be able to predict the impact of a new feedback item on seemingly unrelated target component weights, that impact can be justified and (if desired) traced to specific judgments. Again, the tradeoff is largely a function of complexity of representation, and a more complex Bayesian model would most likely provide the same benefits and encounter the same costs of complexity as the present fuzzy-sets approach.

4.4.3 Use of computational resources - This issue refers to the demands of each method on computing time, complexity, and storage space. In this area, given the highly constrained resources of the system on which the prototype targeting aid has been implemented, the Bayesian approach offers a distinct advantage over the fuzzy-sets method. This advantage stems from two distinct properties of the Bayesian approach chosen: the relative simplicity of an N-parameter representation; and the selection of a family of distributions which permits very simple arithmetical operations in a non-iterative analytic updating algorithm.

By contrast, the fuzzy-sets method described in this report involves an N(N-1)-parameter model representation, and requires a relatively complex updating procedure which searches in an iterative or recursive manner for critical discrepant cycles. Moreover, the selection of a critical cycle involves the calculation of n^{th} roots, a process which would be very inefficient in the current implementation.

The Bayesian approach can be successfully implemented using an N x 3 array to represent the prior, adjusted, and updated weights (in the present problem, N will be approximately 15). For the same problem, the current version of the fuzzy-sets algorithm requires the construction of an N x N x N array to represent the partial cycles of various lengths, as well as the N x N matrix which represents the inequalities. (Practically, the N x N x N matrix could be reduced to an N x N matrix, but that matrix would have to be constructed N times; in a space-constrained environment this could be accomplished, but only at a cost in terms of speed.)

Because the fuzzy-sets algorithm could not be implemented on the targeting aid using the current system configuration, no valid real-time comparison can be made at present. It seems clear that given the relative complexities of the two systems, the current Bayesian procedure would almost always run faster than the fuzzy-sets routine; however, the demonstrated ability of the IBM 4331 implementation of the fuzzy-sets method to produce an answer in under 20 seconds might indicate that either approach would be acceptable if implemented on the proper system.

5.0 CONCLUSIONS AND RECOMMENDATIONS

The major goal of this project was to compare the theory of fuzzy subsets with Bayesian theory in terms of their usefulness for decision aids. In order to accomplish this goal, it was necessary to create a context for a decision aid which was consistent with the spirit of theories of approximate reasoning. Much of the effort in this project was concerned with creation of a basic framework for the aid which was consistent with the notion that human judgment is often imprecise and approximate. In order to develop this framework, considerable effort was devoted to developing graphical-display, assessment, and analytical methods.

Specific goals of the project included the following:

1. Development of a meaningful framework for a target nomination aid.

2. Development and implementation of the capability for natural display of maps, and overlay of targeting data on the map background.

3. Development of aid components based on the theory of fuzzy subsets and on Bayesian methods and evaluation of performance of these aid components.

In order to accomplish the goals listed above, it was necessary to set aside some of the problems needing to be solved if a decision aid for actual use were being developed. These considerations include:

1. Development and maintenance of the target data base.

2. Development of specific models of target capacity/ importance.

3. Inclusion of mission planning factors and some other factors relevant to the analysis.

This report has covered three aspects of the targeting aid: the map display system, the target nomination analysis, and the feedback adjustment procedures. The conclusions and recommendations will be organized according to this division of aid functions.

5.1 Geographic Display System

The geographic display system provides a rich context for the display of targeting information and results, and gives the targeting analyst the capability for natural interaction with maps covering a wide area. In addition, the user has considerable flexibility to construct a map which displays the features that are most interesting for the task at hand.

Work in the geographic display system has uncovered four areas in which considerable improvements could be made in the map display capabilities. These areas involve storage requirements for map data, human factors issues, image quality issues, and alternate data capture methods.

5.1.1 Storage requirements - Even though the videodisc provides the capability to store a great number of map images, it is still important to develop methods which conserve use of videodisc storage. Conservation is required because of requirements for wide coverage and great detail, and because of the limited resolution of the display system which requires single frames to cover only a small segment of a map sheet.

The current geographic display system requires a large amount of videodisc storage for two reasons. First, maps are stored as up to twelve map separates occupying separate frames on the videodisc. Second, the maps were photographed at 66 percent overlap. This high level of overlap requires nine times the storage allocation that would be required if maps could be photographed with no overlap.

There are ways of improving both of these problems, but additional research is necessary to demonstrate their feasibility. With regard to the number of separates, it may be possible to combine the information from several map separates on a single videodisc frame. Since feature separates contain one, or at most two bits of information per pixel, the combination of several separates on a single frame is at least feasible. The major determinant of how many separates can be combined is the number of bits of information that can reliably be recovered from the videodisc. A current guess is that four bits should be recoverable. If this guess is accurate, up to four separates may be combined. In addition, for some applications, the control of the display may not justify the space and image processing capabilities required to process map separates, and a simpler, more efficient method may be preferred.

The high degree of overlap is dictated to some extent by the requirement for continuous scrolling. The method which is used for scrolling requires a minimum of 50 percent overlap. The additional overlap was used to give smoother scrolling in situations in which the user changes directions suddenly. If this capability is given up, a reduction to 50 percent overlap would approximately halve the storage requirements. If discrete scrolling is deemed adequate, evidence indicates that 25 percent overlap may be adequate. A final potential solution to the problem of storage requirements, brought about by continuous scrolling, is to change the method of scrolling. One potential scrolling method would use the image processor to piece together the to-be-displayed image from several videodisc frames. This method may not be possible with the current image-processing hardware (or it may require relinquishing the capability of overlay of targeting data). However, if it is feasible, the storage requirements would be reduced substantially, and continuous scrolling would be maintained.

5.1.2 Human factors issues - There are several issues that were discussed in Section 2.0 that relate to the ability of the user to interact naturally with the map display system. These issues will be mentioned here briefly. First, a number of users tended to become disoriented when they zoomed to a larger or smaller scale map. Although this tendency decreased as the user became more accustomed to the system, there may be ways to reduce or eliminate the problem. One alternative is continuous zooming. This alternative may not be feasible with the current hardware; however it offers an attractive alternative which may help dictate the choice of hardware for a future system. A second alternative would be to display a window indicating the area that will be covered if the user zooms. Under this simple alternative, the user may adjust the size and location of the window so that it covers the desired area and zooms directly to the area covered by the window. Other alternatives, which use the touchscreen as an input device, may be feasible as well.

Other human factors issues relate to the techniques used to overlay target information on the map background. It is probably desirable to aggregate targets into clusters for display on small scale maps. Aggregation could be accomplished by developing a hierarchical organization of targets: at small scales, only aggregated targets would be displayed;

at more detailed scales, individual targets could be displayed. In addition, the methods used to display target attributes using color should take advantage of the abundant research in this area. However, there are many pitfalls in the display of data using color; these should be avoided in designing decision aids.

5.1.3 Image quality - The major issues in image quality involve the processing of map separates to produce an integrated image. The separates essentially represent digital information; yet, it is often difficult to recover this information because of noise introduced by the processes of transferring the information to the videodisc and recovering the information in the frame buffer. To solve this problem it is necessary to develop some context-sensitive threshold scheme. Such a scheme should increase the distinction between signal and background on the separates. In order to maintain rapid access to maps, it may be necessary to perform image enhancement off-line, producing a second videodisc containing refined images.

Other problems in image quality include shimmering, moire patterns, and aliasing. Some of these problems may also be improved by off-line processing using recent techniques of computer graphics (see Whitted, 1982). Others may require processing in real time; the benefits of using these techniques would need to be compared against the costs in access time. An alternate solution to the problems of image quality involves use of some method other than map separates. The use of map separates affects many aspects of the display system, and the benefits should be weighed against the cost for any specific application.

5.1.4 Alternative data capture - The process by which data were transferred to the videodisc involves several steps in which the maps were taken from the original source, acetate sheets, and placed on film, videotape, and finally videodisc. Alternate data capture methods may be beneficial, but little research has been performed on this subject. Sources of map data include paper maps and digital terrain models, in addition to map separates. Storage media include optical disks and a digital representation on a videodisc in addition to an analog representation on a videodisc. There are potential advantages to each of the alternate data sources and storage media. Development and evaluation of methods using these alternatives should be pursued.

5.2 Target Nomination Analysis

The target nomination analysis is a simple, decision-analytic model which is designed to capture the gist of a problem. However, because of the emphasis on approximate reasoning, it was impossible to expend sufficient effort on tailoring the analysis to the requirements of the specific user. In addition, certain factors were not considered in the analysis, and no formal evaluation was conducted. These three issues are addressed in the following subsections.

5.2.1 User involvement in aid development - Research has indicated that decision aids are most successful when the user is intimately involved in their development (Adelman, Donnell, Patterson, and Weiss; 1981). However, because this project was principally concerned with development of alternate methods for analysis and display of targeting data, and because the methods have general applicability for a variety of tasks, only minimal efforts were made to tailor the aid to a specific user. Now that the methods have been investigated, it is necessary to develop an aid for a specific user, probably the targeting officer. In developing the specific aid, there will undoubtedly be major changes in the model from the demonstration aid developed in this effort. Objectives will probably be changed, and some will be eliminated. Certain targets will become irrelevant, although it may be necessary to display them even though they do not enter into the analysis. The eventual form of the analysis will be determined through intensive interactions between the model developer and the user, although the concepts developed in this effort will provide a valuable framework for target nomination analysis.

5.2.2 Additional factors in target nomination - Several critical factors in the design and operation of the decision aid were tabled for consideration at a later time because of the emphasis of the project. These issues include development and maintenance of the target data base, development of specific models of target capacity/importance, and inclusion of additional factors in the analysis. Further research and development efforts should concentrate on these issues, so that a practical, accurate, and efficient aid is developed. The aid should be consistent with the constraints of current capabilities in the area of data management, as well as with the needs of the user.

5.2.3 Evaluation of the analysis - The eventual aid will provide a great change in the way that target nomination is carried out. It is important that the effects of this change are carefully evaluated so that the aid leads to an efficient and effective target nomination process. The eventual aid should be formally evaluated by experimental or quasi-experimental methods. The effort and resources allocated to formal evaluation will increase the likelihood that the lessons learned in the development of the aid will be transferred to future efforts.

5.3 Feedback Adjustment

The goal of the feedback-adjustment method using fuzzy subsets was to provide a way in which a decision aid may make sophisticated inferences from simple and natural input. Although the adjustment mechanism encountered certain system-related difficulties, the goal of sophisticated adjustment is still valid. Furthermore, the current fuzzy-sets algorithm is a new one, and several improvements and extensions may be possible. The Bayesian method does not provide the sophistication desired in this area.

It is probably necessary for research in this area to consider alternatives to the feedback language provided in the current aid. This language should provide the user a richer capability to express criticism of the model results, while at the same time be natural to the user. If an appropriate language for user feedback is developed, then there will be a variety of methods available to process the feedback and adjust model parameters accordingly. Both of the candidate methods described in this report would be available for modification to the new feedback language, as will other methods.

However, both methods are expected to have some problems with a richer representation of feedback. The problem with the Bayesian methods is to find conjugate prior distributions which do not produce an overly simple representation of the feedback. The fuzzy-subset method will have to develop a mechanism to use the richness of the feedback while avoiding the problems that it had in the current system. Although the Bayesian method is the method of choice for the current representation of feedback, the choice is unclear, given an expanded feedback language. In any event, future research in this area will concentrate more on how the user should be allowed to express criticism and how this expression of criticism should be processed by the aid, and less on a straightforward comparison between two specific methods.

REFERENCES

Adelman, L.; Donnell, M. L.; Patterson, J. F.; and Weiss, J. J. Issues in the Design and Evaluation of Decision-Analytic Aids. Technical Report TR 81-1-304. McLean, VA: Decisions and Designs, Inc., February 1981.

Allen, J. J.; Kelly, C. W., III; Phillips, L. H.; and Stewart, R. R. Computer-Assisted Option Screening and Intelligence Assessment: Software Users' Guide. McLean, VA: Decisions and Designs, Inc., October 1976.

Anderson, R. H., and Shapiro, N. Z. Design Considerations for Computer-Based Interactive Map Display Systems. Final Report R-2382-ARPA. Santa Monica, CA: Rand Corp., February 1979.

Barclay, S; Esoda, R. M.; Cox, C. L.; Peterson, C. R.; and Weiss, J. J. A Friendly Personal Decision Aid. Technical Report TR 81-12-312.43. McLean, VA: Decisions and Designs, Inc., December 1981.

DeGroot, M. H. Optimal Statistical Decisions. New York: McGraw-Hill, 1970.

Goldstein, C. M. "Optical Disk Technology and Information." Science 215 (1982):862-868.

Humphreys, P. C., and Wisudha, A. MAUD - An Interactive Computer for the Structuring, Decomposition, and Recomposition of Preference Between Multiattributed Alternatives. Decision Analysis Unit Technical Report 79-2. Uxbridge, Middlesex, England: Brunel University, 1979.

Kelly, C. W., III; Gulick, R. M.; and Stewart, R. R. The Decision Template Concept. Final Technical Report PR 80-17-99. McLean, VA: Decisions and Designs, Inc., September 1980.

Kibler, A. W.; Watson, S. R.; Kelly, C. W., III; and Phelps, R. H. A Prototype Aid for Evaluating Alternative Courses of Action for Tactical Engagement. Final Technical Report 78-38A. Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, November 1978.

Kuskey, K. P. "Advanced Network-Targeting Aid: System Design Concepts." Working paper prepared for the Defense Communications Agency and Rapid Deployment Joint Task Force, McLean, VA, April 1981.

Leal, A., and Pearl, J. A Computer System for Conversational Elicitation of Problem Structures. Research Report UCLA-ENG-REP-7665. Los Angeles: University of California, June 1976.

Merkhoffer, M. W.; Robinson, B.; and Korsan, R. J. A Computer-Aided Decision Structuring Process. Menlo Park, CA: SRI International, June 1979.

Moses, F. L., and Maisano, R. E. User Performance Under Several Automated Approaches to Changing Displayed Maps. Technical Paper 366. Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, June 1979.

Negroponte, N. Spatial Data-Management. Cambridge, MA: Massachusetts Institute of Technology, 1979.

Patterson, J. F.; Randall, L. S.; and Stewart, R. R. Advisory Decision Aids: A Prototype. Technical Report PR 80-27-312. McLean, VA: Decisions and Designs, Inc., February 1981.

U. S. Air Force. An Introduction to Air Force Targeting. AF Pamphlet 200-17, October 1978.

Weiss, J. J. QVAL and GENTREE: Two Approaches to Problem Structuring in Decision Aids. Technical Report TR 80-3-97. McLean, VA: Decisions and Designs, Inc., April 1980.

Weiss, J. J., and Kelly, C. W., III RSCREEN and OPGEN: Two Problem Structuring Decision Aids Which Employ Decision Templates. Technical Report TR 80-4-97. McLean, VA: Decisions and Designs, Inc., October 1980.

Whitted, T. "Some Recent Advances in Computer Graphics." Science 215 (1982):767-774.

Zadeh, L. A. "Fuzzy Sets." Information and Control 8 (1965): 338-353.

Appendix A

HARDWARE CONFIGURATION

The hardware for the target nomination aid is depicted in Figure A-1. It consists of two separate but interacting subsystems. One is designed to provide rapid access, manipulation, and display of whole image data (512 x 512 bytes). The second is a standard digital computer, which implements the non-image processing and acts as a host for the first subsystem.

The most distinctive aspect of the hardware is the whole image processing subsystem. Images that have been stored on the videodisc can be retrieved and placed in one of the DeAnza's four image memories. Since the image presented by the videodisc player is in an NTSC <u>analog</u> format, it must be digitized before it can be stored in one of the DeAnza digital image memories. The analog-to-digital converter is a feature of the DeAnza unit itself. Once stored in an image memory, an image can be scrolled, zoomed, overlayed, and operated on in a variety of ways. Finally, the image is displayed on the Ramtek color CRT. (At this stage the DeAnza unit converts the digital image back to an appropriate analog format.)

The time-based corrector is needed to stablize the output of the videdisc player for the DeAnza image processor. Because of signal fluctuations introduced by variations in the videodisc player's rotational speed a given image will "wobble" over time. "Snapshots" of that image taken by the DeAnza will tend to be slightly distorted, causing feature overlays to be out of registration. The time-based corrector compensates for these fluctations.

Besides operating on whole images, the system must also manipulate digital data. A Digital Equipment Corporation PDP 11/40 performs this function. This system has two RK05 and two RL02 magnetic disk drives for peripheral storage, as well as an alphanumeric CRT, a keyboard, an Elographics Touchscreen and a DeAnza Systems joystick with special function keys.

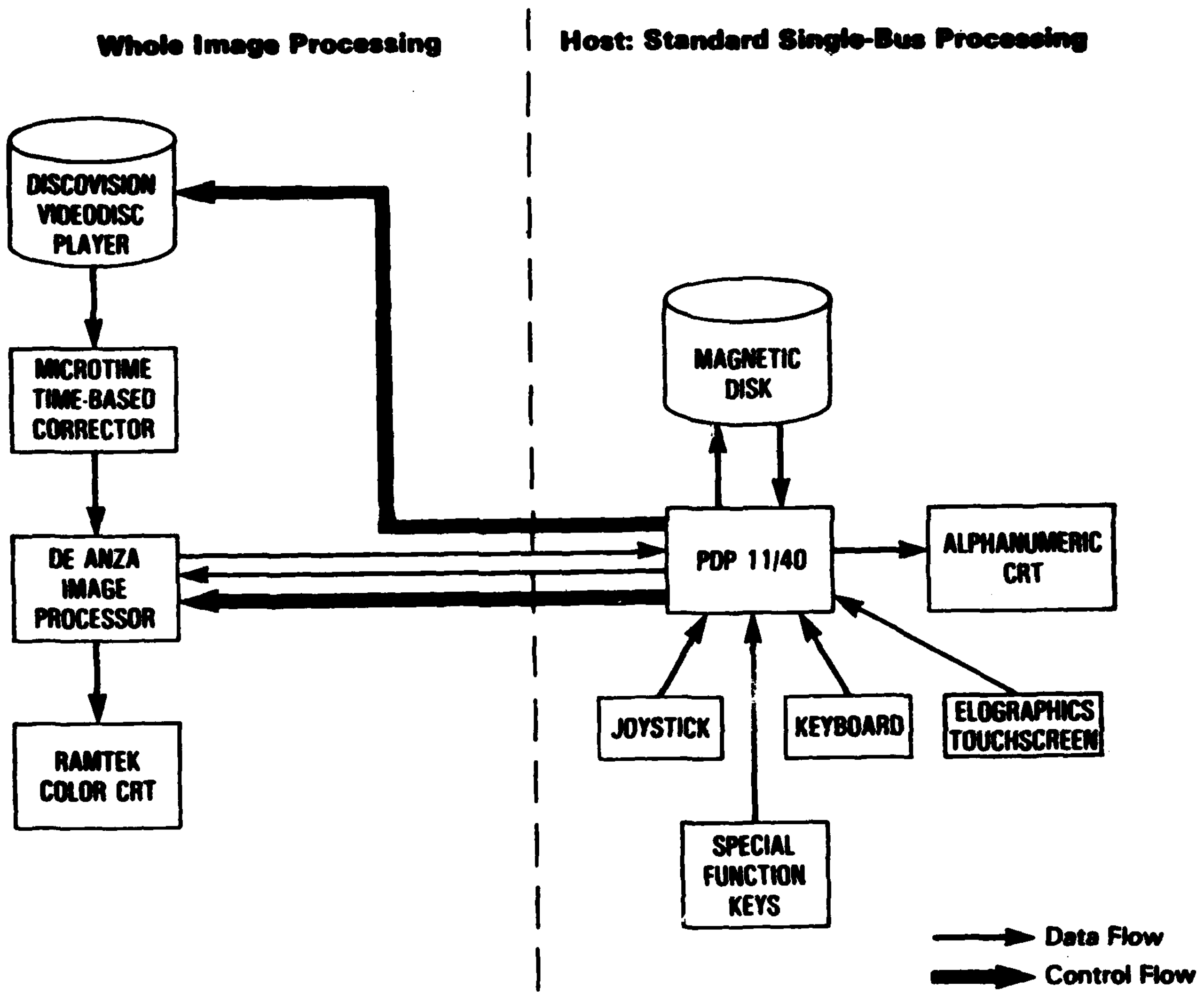

Figure A-1
HARDWARE CONFIGURATION

In addition to performing digital calculations, the PDP 11/40 serves as a controller for the image processor and videodisc. Thus, while the image processing and digital processing are implemented by separate devices, they are controlled by a single device.

A final feature of the hardware configuration is a facility to transmit information between the image processor memories and the PDP 11/40 memory. This capability allows portions of images to be modified as a function of information in the host's memory. In particular, it is used to overlay target symbols.

The key to understanding this design is to appreciate that the maps are used primarily as a static backdrop. They are not analyzed or processed except in the sense that map feature separates are combined. All dynamic information is overlayed by the PDP 11. The computer does not "know" about the information contained in the map. It simply uses the maps to establish a context for the data that it presents.